

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2002, 03:48:42 ; Search time 25 seconds

(without alignments)
535.879 Million cell updates/sec

Title: US-09-886-041-2

Perfect score: 1853

Sequence: 1 MYNGSCRIEQTISQVMP.....ANSFQSDQWDPHIVEWH 346

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	880.5	47.5	387	1 HM74_HUMAN	P49019 homo sapien
2	448.5	24.2	319	1 GPRV_HUMAN	O00270 homo sapien
3	371.5	20.0	309	1 CLT2_MOUSE	Q920a1 mus musculu
4	366.5	19.8	346	1 CLT2_HUMAN	Q9ns75 homo sapien
5	364	19.6	362	1 P2YR_CHICK	P34996 gallus gall
6	364	19.6	362	1 P2YR_MELGA	P49852 melagris g
7	362.5	19.6	373	1 P2YR_BOVIN	P48042 bos taurus
8	362	19.5	345	1 CLT2_PIG	Q95n03 sus scrofa
9	358.5	19.3	373	1 P2YR_RAT	P49651 rattus norv
10	355.5	19.2	309	1 CLT2_RAT	Q924t9 rattus norv
11	354	19.1	373	1 P2YR_HUMAN	P47900 homo sapien
12	351.5	19.0	373	1 P2YR_MOUSE	P49650 mus musculu
13	346.5	18.7	370	1 P2Y9_HUMAN	Q99677 homo sapien
14	339	18.3	308	1 P2Y5_CHICK	P32250 gallus gall
15	336	18.1	339	1 GPRH_HUMAN	Q13304 homo sapien
16	327	17.6	309	1 GPR2_HUMAN	Q9hc97 homo sapien
17	324	17.5	344	1 P2Y5_HUMAN	P43657 homo sapien
18	320	17.3	373	1 P2UR_MOUSE	P35383 mus musculu
19	316.5	17.1	374	1 P2UR_RAT	P41232 rattus norv
20	310.5	16.8	420	1 THRR_XENLA	P47749 xenopus lae
21	309.5	16.7	363	1 SSR5_RAT	P30938 rattus norv
22	308.5	16.6	307	1 GPR2_MOUSE	Q9es90 mus musculu
23	308	16.6	340	1 CLT1_PIG	Q95n02 sus scrofa
24	307	16.6	381	1 SSR1_MOUSE	P30873 mus musculu
25	305.5	16.5	391	1 SSR1_HUMAN	P30872 homo sapien
26	305.5	16.5	391	1 SSR1_RAT	P28646 rattus norv
27	305	16.5	337	1 CLT1_HUMAN	Q9y271 homo sapien
28	305	16.5	377	1 P2UR_HUMAN	P41231 homo sapien
29	304.5	16.4	359	1 AG2R_BOVIN	P25104 bos taurus
30	302.5	16.3	369	1 SSR2_HUMAN	P30874 homo sapien
31	302	16.3	391	1 BRB2_HUMAN	P30411 homo sapien
32	302	16.3	399	1 PAR2_MOUSE	P55086 mus musculu
33	300.5	16.2	369	1 SSR2_MOUSE	P30875 mus musculu

34	300.5	16.2	369	1 SSR2_RAT	P30680 rattus norv
35	300	16.2	328	1 P2Y6_RAT	Q63371 rattus norv
36	300	16.2	384	1 SSR4_RAT	P30937 rattus norv
37	300	16.2	428	1 SSR3_MOUSE	P30935 mus musculu
38	299.5	16.2	369	1 SSR2_PIG	P34994 sus scrofa
39	299	16.1	388	1 SSR4_HUMAN	P31391 homo sapien
40	299	16.1	397	1 PAR2_RAT	Q63645 rattus norv
41	298.5	16.1	418	1 SSR3_HUMAN	P32745 homo sapien
42	298.5	16.1	432	1 THRR_RAT	P26834 rattus norv
43	298	16.1	375	1 GPRX_MOUSE	Q9rlk6 mus musculu
44	298	16.1	428	1 GPR3_RAT	P30936 rattus norv
45	297	16.0	397	1 PAR2_HUMAN	P55085 homo sapien

ALIGNMENTS

RESULT 1	HM74_HUMAN	STANDARD;	PRT;	387 AA.
ID	HM74_HUMAN			
AC	P49019;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Probable G protein-coupled receptor HM74.			
GN	HM74.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Monocytes;			
RX	MEDLINE=94052629; PubMed=7505609;			
RA	Nomura H., Nielsen B.W., Matsushima K.;			
RT	"Molecular cloning of cDNAs encoding a LD78 receptor and putative leukocyte chemotactic peptide receptors.";			
RL	Int. Immunol. 5:1239-1249(1993).			
CC	-1- FUNCTION: ORPHAN RECEPTOR.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; D10923; BAA01721.1; -			
DR	HSSP; P34996; 1DDD.			
DR	GCRDB; GCR_0528; -			
DR	MIM; 606039; -			
DR	InterPro; IPR000276; GPCR_Rhodpsn.			
DR	Pfam; PF00001; 7tm_1; 1.			
DR	PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.			
DR	PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.			
KW	G-protein coupled receptor; Transmembrane.			
FT	DOMAIN 1 28			
FT	TRANSMEM 29 50			
FT	DOMAIN 51 63			
FT	TRANSMEM 64 85			
FT	DOMAIN 86 102			
FT	TRANSMEM 103 123			
FT	DOMAIN 124 142			
FT	TRANSMEM 143 163			
FT	DOMAIN 164 194			
FT	TRANSMEM 195 209			
FT	DOMAIN 210 236			
FT	TRANSMEM 237 256			
FT	DOMAIN 257 273			
FT	TRANSMEM 274 298			

FT DOMAIN 299 387 CYTOPLASMIC (POTENTIAL).
FT DISULFID 100 177 BY SIMILARITY.
SQ SEQUENCE 387 AA; 44481 MW; C244F562C2343647 CRC64;

Query Match 47.58; Score 880.5; DB 1; Length 387;
Best Local Similarity 52.28; Pred. No. 1.1e-54;
Matches 178; Conservative 49; Mismatches 107; Indels 7; Gaps 4;

Qy 5 CRIEGDTISQVMPPLLIIVAFVILGALNGVALCGFCFHMKTWKPSTVYLFENLAVADFL 64
Db 17 NCCVPRDDFIKAPVPLVGLFIFGLLNGLALWFCFHLKSWKSRIFENLAVADFL 76
Qy 65 MCLPFRDYLLRRHWAFGDIQPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVHPHRAV 124
Db 77 IICLPFVMDYVRRSDNMGDIPCLRLVLFMEAMNRQSGIIFLTVVAVDRYFVHPHRAH 136
Qy 125 NTISRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCSFIMESANGWHDMFOLE 184
Db 137 NKISNWTAAIISCLMGITVGLTVHLLKKLLIQNGPANVCISFSICHTFRWHEAMFLE 196
Qy 185 FFMPLGIILFCSEFKIVSLRRQOLARQARKKATFTIMVAIVFITCYLPSSVSARLYFL 244
Db 197 FLPLGIILFCSEKARILNSLRQ-QMDRAKIKRAITFTIMVAIVFVFCFLPSVVVRIRIF 255
Qy 245 WTVFSSA---CD--PSVHGALHITLSFTYMNMSMLDPLVYFSSFPKFKYKLCISLKP 299
Db 256 WLLHTSGTQNCVRSVDLAFITLSFTYMNMSMLDPLVYFSSFPKFKYKLCISLKP 315
Qy 300 KOPGHKTRQPEEMPISNLGRSICISVANSFQSQSDGOWDP 340
Db 316 KWTGPDNNRSTVELTGDPNKT-RGAPEALMANSGEPMSP 355

RESULT 2
GPRV_HUMAN STANDARD; PRT; 319 AA.
AC 000270;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable G protein-coupled receptor GPR31.
GN GPR31.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=97349123; PubMed=9205127;
RA Zingoni A., Rocchi M., Storlazzi C.T., Bernardini G., Santoni A.,
RA Napolitano M.;
RT Isolation and chromosomal localization of GPR31, a human gene
RT encoding a putative G protein-coupled receptor.";
RL Genomics 42:519-523(1997).
CC -1- FUNCTION: ORPHAN RECEPTOR.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U65402; AAC51375.1; -
CC GCRdb; GCR_1331; -
CC MIM; 602043; -
CC InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tmL1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.

DR PROSITE; PS00237; G_PROTEIN_RECEP_FL1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_FL2; 1.
KW G-protein coupled receptor; Transmembrane.
FT DOMAIN 1 16 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 17 37 1 (POTENTIAL).
FT DOMAIN 38 52 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 53 73 2 (POTENTIAL).
FT DOMAIN 74 91 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 92 110 3 (POTENTIAL).
FT DOMAIN 111 131 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 132 152 4 (POTENTIAL).
FT DOMAIN 153 180 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 181 201 5 (POTENTIAL).
FT DOMAIN 202 219 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 220 240 6 (POTENTIAL).
FT DOMAIN 241 265 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 266 284 7 (POTENTIAL).
FT DOMAIN 285 319 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 5 N-LINKED (GLCNAC..) (POTENTIAL).
SQ SEQUENCE 319 AA; 35079 MW; 7A4F164CD9C969DD CRC64;

Query Match 24.28; Score 448.5; DB 1; Length 319;
Best Local Similarity 34.08; Pred. No. 1.1e-24;
Matches 100; Conservative 62; Mismatches 121; Indels 11; Gaps 5;

Qy 7 CRIEGDTISQVMPPLLIIVAFVILGALNGVALCGFCFHMKTWKPSTVYLFENLAVADFLMI 66
Db 6 CSAFTVAVTAGVLLGLECGLLGNAVALWTFTRVVRVMPKPYAVYLLNALADLLAA 65
Qy 67 CLPFRDYLLRRHWAFGDIQPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVHPHRAVNT 126
Db 66 CLPFAAFYLSLQAHHLGRVGCWALRFLDLRSRSGMAFLAVALDRYLVVHPKLYNL 125
Qy 127 ISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCSFIMESANG-----WHDIMF 181
Db 126 LSPQALGSLVLLWLLVALTCPGLLISE--AAQNSTCHSF-YSRADGSSIIHQEALS 182
Qy 182 QLEFFMPLGIILFCSEFKIVSLRRR-QOLARQARKKATFTIMVAIVFITCYLPSSVSAR 240
Db 183 CLQFVLPGLVFCNAGIIRALQKRLREPKQPKLQRAQALVTLVVLVFCFLPCFLAR 242
Qy 241 --LYFLWTVPSACDPSVHGALHITLSFTYMNMSMLDPLVYFSSFPKFKYKNL 292
Db 243 VLMHIFQNLGSCALCAVAHTSDVTGSLTYLHVVVNPVYVFCSSPTFRSRRV 296

RESULT 3
CLT2_MOUSE STANDARD; PRT; 309 AA.
ID CLT2_MOUSE
AC Q920A1;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Cysteinyl leukotriene receptor 2 (CysLTR2).
GN CysLTR2 OR CysLTR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX PubMed=11591709;
RA Hui Y., Yang G., Galczenski H., Figueroa D.J., Austin C.P.,
RA Copeland N.G., Gilbert D.J., Jenkins N.A., Funk C.D.;
RA "The murine cysteinyl leukotriene 2 (CysLTR2) receptor, cDNA and
RT genomic cloning, alternative splicing, and in vitro
RT characterization.";
CC J. Biol. Chem. 276:47489-47495(2001).
CC -1- FUNCTION: Receptor for cysteinyl leukotrienes. The response is
CC mediated via a G-protein that activates a phosphatidylinositol-
CC calcium second messenger system. The rank order of affinities for
CC the leukotrienes is LTC4 > LTD4 >> LTE4.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Widely expressed at low levels, with highest
CC expression in the spleen, thymus and adrenal gland, and lower in
CC the kidney, brain and peripheral blood leukocytes.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF331658; AAK97354.1; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm.1; 1.
DR PROSITE: PS00237; G_PROTEIN_RECIP_FL1; FALSE_NEG.
DR PROSITE: PS00262; G_PROTEIN_RECIP_FL2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 26 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 27 47 1 (POTENTIAL).
FT DOMAIN 48 56 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 57 77 2 (POTENTIAL).
FT DOMAIN 78 98 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 99 119 3 (POTENTIAL).
FT DOMAIN 120 138 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 139 159 4 (POTENTIAL).
FT DOMAIN 160 187 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 188 208 5 (POTENTIAL).
FT DOMAIN 209 229 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 230 250 6 (POTENTIAL).
FT DOMAIN 251 271 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 272 292 7 (POTENTIAL).
FT DOMAIN 293 309 CYTOPLASMIC (POTENTIAL).
FT DISULFID 95 171 BY SIMILARITY.
FT CARBOHYD 14 14 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 309 AA; 35227 MW; 327B14A6EDDD2A02 CRC64;

Query Match 20.0%; Score 371.5; DB 1; Length 309;
Best Local Similarity 32.7%; Pred. No. 2.5e-19;
Matches 96; Conservative 56; Mismatches 125; Indels 17; Gaps 7;

Qy 2 YNGSCCRIGETISQVMPPLIIVFALGALNGVALGFCFHMKTWKPST---VYLFMLA 58
Db 1 YSNRNCTIE-NFKKEFYPIIYLIIFWALGNGFSIYF---LQCKKSTSVNFMNL 65
Qy 59 VADFLMLICLPFDYDYLRRHAFGDIPICRVGLFTLANNRAGSIYFLTVVVAADRYFKV 118
Db 66 TSDFLFTLFPADYFPRGSNFIWFGDLACRVMSYSLYVNYMTSYIYFLTVLSVVRFOATV 125
Qy 119 HPHAVNTISTRVAGIIVCTLWALVILGTVYLLLENHLCVQETAVSC-----ESFIMESA 173
Db 126 HPRFMHVTSVR-SANILCGIIVFWIMASSALLVNGQEKNDIISCULESPQKF--KSL 182
Qy 174 NGWHDIMFQLEFFPLGLIILFCSEKIYMSLRRO--QLARQARKMKATFRIMVAIVFIT 231
Db 183 LIMNHIAVAGFLPLFTLTVCYLLIIRILKAEIPESGPRAHRKALTIVIAMITLL 242
Qy 232 CYLPSVSARLYFLWTVPSSACDPSVHGALHTLSFTYKNSMLDPLVYFSPSP 285
Db 243 CFLPYHALRTLHLVTWDRKDCGDVLHKTATVITLMAAANSFNFPFLYFAGENF 296

RESULT 4
CLT2_HUMAN
ID CLT2_HUMAN STANDARD; PRT; 346 AA.
AC Q9NS75; Q9HCQ2;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Cysteinyl leukotriene receptor 2 (CysLTR2) (PSEC0146) (HG57) (HPN321).
GN CysLTR2 OR CysLTR2 OR CysLTR2.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=20374466; PubMed=10913337;
RA Takasaki J., Kamohara M., Matsumoto M., Saito T., Sugimoto T., Ota T.,
RA Nishikawa T., Kawai Y., Masuho Y., Isogai T., Suzuki Y., Sugano S.,
RA Furuchi K.;
RT "The molecular characterization and tissue distribution of the human
RT cysteinyl leukotriene CysLT2 receptor.";
RL Biochem. Biophys. Res. Commun. 274:316-322(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20459128; PubMed=10851239;
RA Heise C.E., O'Dowd B.F., Figueroa D.J., Sawyer N., Nguyen T.,
RA Im D.-S., Stocco R., Bellefeuille J.N., Abramovitz M., Cheng R.,
RA Williams D.L., Jr., Zeng Z., Liu Q., Ma L., Clements M.K., Coulombe N.,
RA Liu Y., Austin C.P., George S.R., O'Neill G.P., Metters K.M.,
RA Lynch K.R., Evans J.F.;
RT "Characterization of the human cysteinyl leukotriene 2 receptor.";
RL J. Biol. Chem. 275:30531-30536(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX PubMed=11093801;
RA Nothacker H.-P., Wang Z., Zhu Y., Reinscheid R.K., Lin S.H.S.,
RA Civelli O.;
RT "Molecular cloning and characterization of a second human cysteinyl
RT leukotriene receptor: discovery of a subtype selective agonist.";
RL Mol. Pharmacol. 58:1601-1608(2000).
RN [4]
RP SEQUENCE FROM N.A.
RX Dunn M.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 17-346 FROM N.A.
RA Suga H.;
RT "Homo sapiens cysteinyl leukotriene receptor 1 like receptor.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Receptor for cysteinyl leukotrienes. The response is
CC mediated via a G-protein that activates a phosphatidylinositol-
CC calcium second messenger system. Stimulation by BAY u9773, a
CC partial agonist, induces specific contractions of pulmonary veins
CC and might also have an indirect role in the relaxation of the
CC pulmonary vascular endothelium. The rank order of affinities for
CC the leukotrienes is LTC4 > LTD4 > LTE4.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Widely expressed, with highest levels in the
CC heart, placenta, spleen, peripheral blood leukocytes and adrenal
CC gland. In lung, expressed in the interstitial macrophages, and
CC slightly in smooth muscle cells.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB038269; BAB03601.1; -
DR EMBL: AF254664; AAG17281.1; -
DR EMBL: AF279611; AAK69485.1; -
DR EMBL: ALI37118; CAC29102.1; -
DR EMBL: AB041644; BAB16379.1; -
DR MIM: 605666; -
DR InterPro: IPR004071; Cysleuk_receptor.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm.1; 1.
DR PRINTS: PR01533; CYSLTRECPTP.

```
DR PRINTS: PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; FALSE_NEG.
DR PROSITE; PS0262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 42
FT TRANSMEM 43 63
FT DOMAIN 64 72
FT TRANSMEM 73 93
FT DOMAIN 94 123
FT TRANSMEM 124 144
FT DOMAIN 145 153
FT TRANSMEM 154 174
FT DOMAIN 175 204
FT TRANSMEM 205 225
FT DOMAIN 226 245
FT TRANSMEM 246 266
FT DOMAIN 267 286
FT TRANSMEM 287 307
FT DOMAIN 308 346
FT DISULFID 111 187
FT CARBOHYD 20 26
FT CARBOHYD 26 30
FT CARBOHYD 30 30
FT CARBOHYD 181 181
SQ SEQUENCE 346 AA; 39635 MW; EBS44A42DDCE5EE4 CRC64;

Query Match
Best Local Similarity 19.8%; Score 366.5; DB 1; Length 346;
Matches 94; Conservative 69; Mismatches 134; Indels 17; Gaps 8;

QY 3 NSGCCRIEGDTSQWMPPLIVAFVLGALNGVALCGCFHMKTWKPS---VYLFNLAV 59
DB 27 NSRNCITIE-NFREFFPIVYLIIFFVGLNGLSIYVF---LQPKKSTSVNVFMLNLA 82
QY 60 ADFLLMICLPERTDYLLRRHWAFCGICPVCGLFTLANNRAGSIVFLTVVAADRYEKVVH 119
DB 83 SOLLISTLPFRADYILRGSNWIFGLACRIMSYSLYVNWYSYFLTVLSVFLAMVH 142
QY 120 PHHAVNTISTRVAAGIVCTMALVILGTVYLLLENHLCVQETAVSC---ESFIMESANGW 176
DB 143 PFRLLHVTISR-SAWILCGIILIMASSIMLLDSSGSEQSVTSCLENLKYIAKLQTM 201
QY 177 HDIMFQLEFFMPLGIILFCSFIVNLSRRQ--QLARQARKKATRFIMVAIVITCVL 234
DB 202 NYIALVWGCLLPFFTLISCYLLIIRVLLKVEVPESGLRVSHRKALTITIIIFLCLF 261
QY 235 PVSARLVFLVTPVPSACDPSVHGALHITLSPTYNMSMLDPLVYVFPSPFPKFNKLKI 294
DB 262 PHTHTVHTVTKVGLCKDRUHKALVITLALAAANACFNPLLYTFAGENFK---DRUK- 317
QY 295 CSLKPKQPGHSKTQ 308
DB 318 SALRKGHPQAKTK 331

RESULT 5
ID P2YR_CHICK STANDARD; PRT; 362 AA.
AC P34996;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).
DS P2YR1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauromorpha; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93285340; PubMed=8508924;
```


QY 138 TLMALVIL-----GTVLLLENHLCVQETA-----VSCSEFTMESANGWHD 178
 Db 163 LVWALVAVIAPILFYSGTGVRNRKTTTCYDTTADEYLRSLRVFYVSMCTTVFM----- 214
 QY 179 IMQLEFFMPLGIILFCFSKIVMSLRROQLARQARKKATRTFMVAIVFITCYLP--- 235
 Db 215 -----FCIPFIVILGCGYIVKALYIK-DLONSPLRKSIYLIIVLTFAVSYPFHV 267
 QY 236 ----SVSARLYFLWTPSSACDPSVGHALHTLTSFTWNSMLDPLVYFSSPSPKPYNK 291
 Db 268 MKTLNLRARLDF-OTPMCAFNDKRVYATQVTRGLASINSCVDILYFLAGDTFRRLSR 326
 QY 292 LKICSLKPKQPGHSGKTORPEEMPISNL 318
 Db 327 ATRKSSRRSEP--NVQSKSEMTLNIL 351

RESULT 6
 P2YR_MELGA STANDARD; PRT; 362 AA.
 AC P49652;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 13-JUL-1998 (Rel. 36, Last annotation update)
 DE P2Y purinoreceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor) (G6H1 orphan receptor).
 GN P2Y1.
 OS Meleagris gallopavo (Common turkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Meleagridae; Meleagris. NCBI_TaxID=9103;
 RN [1]
 RP TISSUE=Brain;
 RC MEDLINE=94335907; PubMed=8058061;
 RA Flitz T.M., Li Q., Boyer J.L., Nicholas R.A., Harden T.K.;
 RT "Expression of a cloned P2Y purinergic receptor that couples to phospholipase C.";
 RL Mol. Pharmacol. 46:8-14(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97382436; PubMed=9240460;
 RA Li Q., Schachter J.B., Harden T.K., Nicholas R.A.;
 RT "The G6H1 orphan receptor, claimed to be the P2Y5 receptor, does not mediate nucleotide-promoted second messenger responses.";
 RL Biochem. Biophys. Res. Commun. 236:455-460(1997).
 CC -1- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS ATP AND ADP. SEEMS TO MEDIATE ITS ACTION VIA A PERTUSSIS TOXIN SENSITIVE G-PROTEIN, PROBABLY BELONGING TO THE GQ FAMILY THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: MAINLY FOUND IN BLOOD, BRAIN, AND LUNG. TO A LESSER EXTENT IN STOMACH, GUT, AND SKELETAL MUSCLE.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@lsb-sib.ch).
 CC -----
 DR EMBL; U09842; AAA18784.1; -;
 DR EMBL; AF012103; AAB65428.1; -;
 DR HSSP; P34996; 1DD0.
 DR GCRDB; GCR_2392; -;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCR_RHODOPSIN.
 DR PROSITE; PS00237; G_PROTEIN_RECPEP_FL_1; 1.
 DR PROSITE; PS00262; G_PROTEIN_RECPEP_FL_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.

FT DOMAIN 1 41 EXTRACELLULAR (POTENTIAL).
 FT TRNSKEM 42 63 1 (POTENTIAL).
 FT DOMAIN 64 76 CYTOPLASMIC (POTENTIAL).
 FT TRNSKEM 77 98 2 (POTENTIAL).
 FT DOMAIN 99 115 EXTRACELLULAR (POTENTIAL).
 FT TRNSKEM 116 136 3 (POTENTIAL).
 FT DOMAIN 137 155 CYTOPLASMIC (POTENTIAL).
 FT TRNSKEM 156 177 4 (POTENTIAL).
 FT DOMAIN 178 207 EXTRACELLULAR (POTENTIAL).
 FT TRNSKEM 208 227 5 (POTENTIAL).
 FT DOMAIN 228 254 CYTOPLASMIC (POTENTIAL).
 FT TRNSKEM 255 274 6 (POTENTIAL).
 FT DOMAIN 275 292 EXTRACELLULAR (POTENTIAL).
 FT TRNSKEM 293 317 7 (POTENTIAL).
 FT DOMAIN 318 362 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 113 191 BY SIMILARITY.
 FT CARBOHYD 11 11 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 362 AA; 41180 MW; 3E128AB9EB64349C CRC64;

Query Match 19.6%; Score 364; DB 1; Length 362;
 Best Local Similarity 29.1%; Pred. No. 9.5e-19;
 Matches 95; Conservative 58; Mismatches 130; Indels 44; Gaps 7;

QY 18 MPPLLIIVAFVLGALNGVALCGCFHMKTKPSTVYLFNLAVADFLMLICLPFTDYLR 77
 Db 43 LPTVYILVFTIGFNGSVAIMVFWHMRPWSGIVSYVMFNALADFLVLTLPALIFYFN 102
 QY 78 RRHWAFGDIPCRVGLFTLANNRAGSIVFLTVAAADRFKVVHPHVAIVNTISTRVAAGTVC 137
 Db 103 KTDWIFGDMVKLQRFIFHVNLYGSLFLTCISVHRITGVVHPLKSLGRLKKNAVIVSS 162
 QY 138 TLMALVIL-----GTVLLLENHLCVQETA-----VSCSEFTMESANGWHD 178
 Db 163 LVWALVAVIAPILFYSGTGVRNRKTTTCYDTTADEYLRSLRVFYVSMCTTVFM----- 214
 QY 179 IMQLEFFMPLGIILFCFSKIVMSLRROQLARQARKKATRTFMVAIVFITCYLP--- 235
 Db 215 -----FCIPFIVILGCGYIVKALYIK-DLONSPLRKSIYLIIVLTFAVSYPFHV 267
 QY 236 ----SVSARLYFLWTPSSACDPSVGHALHTLTSFTWNSMLDPLVYFSSPSPKPYNK 291
 Db 268 MKTLNLRARLDF-OTPMCAFNDKRVYATQVTRGLASINSCVDILYFLAGDTFRRLSR 326
 QY 292 LKICSLKPKQPGHSGKTORPEEMPISNL 318
 Db 327 ATRKSSRRSEP--NVQSKSEMTLNIL 351

RESULT 7
 P2YR_BOVIN STANDARD; PRT; 373 AA.
 AC P48042;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE P2Y purinoreceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).
 GN P2RY1.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea; Bovidae; Bovinae; Bos.
 ON NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Aortic endothelium;
 RX MEDLINE=95352058; PubMed=7626079;
 RA Henderson D.J., Elliott D.G., Smith G.M., Webb T.E., Dainty I.A.;
 RT "Cloning and characterisation of a bovine P2Y receptor.";
 RL Biochem. Biophys. Res. Commun. 212:648-656(1995).
 RN [2]

QY 19 PPLLIVAFVGLGNGVALCGFCFHKMTKWPST---VLFNLAVALDFLLMICLPFRDYD 75
 Db 43 PIVLVIFINGALNGFSIIVE---LAPYKSTSVNVFNLALSDLLFTITLPRDYD 99
 QY 76 LRRHWAFGDIPCRVGLFTLAMNRAGSIVLTVVAAADRYKVVHPHNAVNTISTRVAAGI 135
 Db 100 LRGSNXIFGDTPCRMISYVMVNNYSIYFLTVLSVVRFLATVHPFRLHTTSIK-NAWI 158
 QY 136 VCTLWALVILGTVVLLLENHLCVQETAVSC---ESFTMESANGWHDMFOLEFPMPLGII 192
 Db 159 LCGVINIFWASSTVLKNSQKNDVNTLGLNELSNKVTKLWNYVALVGVFVLPGETL 218
 QY 193 LFCFSKIWSLRRRQQLARQM--KKATRFIMVVAIVFITYCLPSYSARLYFLWTVPSS 250
 Db 219 SICVLLIIRALLKVEPESGLRSLHRKALTVIIALIIFELCPUPYHVLRTHLEKNAD 278
 QY 251 ACDSVHGALHITLSFYFMYNSMLDPLVYYPSSPSF 285
 Db 279 KCDRLHKAVAVTLAALAAANSFCNPFLYFAGENF 313

RESULT 9
 P2YR_RAT STANDARD; PRT; 373 AA.
 AC P49651;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DE 01-FEB-1996 (Rel. 33, Last annotation update)
 DE P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).
 GN P2Y1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Insulinoma;
 RX MEDLINE=95298025; PubMed=7779087;
 RA Tokuyama Y., Hara M., Jones E.M.C., Fan Z., Bell G.I.;
 RT "Cloning of rat and mouse P2Y purinoceptors."
 RL Biochem Biophys Res Commun. 211:211-218(1995).
 CC -1- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
 CC ATP AND ADP. IN PANCREATIC ISLETS, MAY MEDIATE SOME OF THE EFFECTS
 CC OF EXTRACELLULAR ATP ON INSULIN SECRETION.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC LUNG, BRAIN, SPLEEN, BUT NOT IN TESTIS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. Usage by and for commercial
 CC or send an email to license@lsb-sib.ch.
 CC
 CC EMBL: U22830; AAA91303.1; -
 DR HSSP; P34996; 1DDD.
 DR GCRdb; GCR_1444; -
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G_PROTEIN_RECP_FL_1; 1.
 DR PROSITE; PS00262; G_PROTEIN_RECP_FL_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 52 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 53 74 1 (POTENTIAL).
 FT DOMAIN 75 87 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 88 109 2 (POTENTIAL).
 FT DOMAIN 110 126 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 127 147 3 (POTENTIAL).
 FT DOMAIN 148 166 CYTOPLASMIC (POTENTIAL).

TRANSMEM 167 188 4 (POTENTIAL).
 FT DOMAIN 189 218 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 219 238 5 (POTENTIAL).
 FT DOMAIN 239 265 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 266 285 6 (POTENTIAL).
 FT DOMAIN 286 303 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 304 328 7 (POTENTIAL).
 FT DOMAIN 329 373 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 124 202 BY SIMILARITY.
 FT CARBOHYD 11 11 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 27 27 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 113 113 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 373 AA; 42321 MW; 6DDF676287B5E648 CRC64;

Query Match 19.3%; Score 358.5; DB 1; Length 373;
 Best Local Similarity 27.0%; Pred. No. 2.3e-18;
 Matches 91; Conservative 68; Mismatches 139; Indels 39; Gaps 8;

QY 18 MPPLIIVAFVGLGNGVALCGFCFHKMTKWPSTVYLFNLAVALDFLLMICLPFRDYDYL 77
 Db 54 LPAYVILVFIIGLNGSVAINMVFHMKPWSGISVYFNLALADFLVLTLPALIFYFN 113
 QY 78 RHHWAFGDIPCRVGLFTLAMNRAGSIVLTVVAAADRYKVVHPHNAVNTISTRVAAGI 137
 Db 114 KTDWIFGDMCKLQRFIFHNLYGSILFLTCISAHRYSGVYVPLKSLGRLKKNAIVSV 173
 QY 138 TLWALVILGTVVLLLENHLCVQET-AVSC-----ESFIMESANGWHDMFOLEFFM 187
 Db 174 LWLVVVAISPIIFYSGTGIRKNKTVCTDSTDEYLSRYFIYSM-----CTTVAMFCI 228
 QY 188 PLGIILFCFSKIWSLRRRQQLARQMKAATRFIMVVAIVFITYCLP-----SVSAR 240
 Db 229 PLVILGCVGLIVRALYK-DLDSNPLRRKSIYIIVLVFAVSYIPFHVWKMKNLRAR 287
 QY 241 LYFLWTVPSACD--PSVHGALHITLSFTYMNMSMLDPLVYYPSSPSKPKFKNKLKICSLK 298
 Db 288 LDF---QTPMCDPNDRVYATYQVTRGLASLNSCVDPLVFLAGDTFRRLSLR----- 337
 QY 299 PKQPGHCKTQRPPEMPISNLGRRSCISVANSFQSQSD 335
 Db 338 ----ATRKASRRSEANLQSKSEMTNLILSEFKONGD 370

RESULT 10
 CLT2_RAT STANDARD; PRT; 309 AA.
 AC Q924T9;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Cysteinyl leukotriene receptor 2 (CysLT2) (RSBP32).
 GN CysLT2 OR CysLT2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kamohara M., Takasaki J., Matsumoto M., Matsumoto S., Saito T.,
 RA Ohishi T., Soga T., Matsushima H., Furuichi K.;
 RT "Characterization of the cloned rat and porcine cysteinyl leukotriene
 RL receptors."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Receptor for cysteinyl leukotrienes. The response is
 CC mediated via a G-protein that activates a phosphatidylinositol-
 CC calcium second messenger system (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```
DR EMBL; AB052661; BAB0816.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECF_F1_1; FALSE_NEG.
DR PROSITE; PS00262; G_PROTEIN_RECF_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 26 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 27 47 1 (POTENTIAL).
FT DOMAIN 48 56 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 57 77 2 (POTENTIAL).
FT DOMAIN 78 98 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 99 119 3 (POTENTIAL).
FT DOMAIN 120 137 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 138 158 4 (POTENTIAL).
FT DOMAIN 159 187 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 188 208 5 (POTENTIAL).
FT DOMAIN 209 229 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 230 250 6 (POTENTIAL).
FT DOMAIN 251 271 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 272 292 7 (POTENTIAL).
FT DOMAIN 293 309 CYTOPLASMIC (POTENTIAL).
FT DISULFID 95 171 BY SIMILARITY.
FT CARBOHYD 14 14 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 167 167 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 309 AA; 35507 MW; A122AC817789D56 CRC64;

Query Local          19.2%; Score 355.5; DB 1; Length 309;
Best Local Similarity 31.2%; Pred. No. 3.2e-18;
Matches 91; Conservative 57; Mismatches 131; Indels 13; Gaps 7;

QY 2 YNGSCRIEGRDTSOMPLLIIVAVLGVAGLNGVACGFCFHMKTWKPST---VYLFNLA 58
Db 10 YSDKNKCTIE-NFKRDFPIIYIIIVGALNGVGFIIYF---LQYKKSTSVNFMFLNA 65
QY 59 VADELLMCLPRTDYILRRRHWAFGDIPCRVGLTFLAMNRAGSIVETLVVAADRYEKV 118
Db 66 ISDFLFTPLPRADYNFRGSDWIEGDMACRIMSVLYNMVTSYIFTLVLSIVRFLATA 125
QY 119 HPHAVNTRISVRAAGIVC-TLWALVILGTVYVLLLENHLCVQETAVSCSEFTMESAN--G 175
Db 126 HPFQMLHITSVR-SAWILCGIIVFIMASSGLLLKHGQEKKNNTTLCFELNLQKFNLVI 184
QY 176 WHIDIMFQLEFFMPLGIILFCSEKIVMSLRRQ--QLARQARKMKATRFIMVVAIVFTCY 233
Db 185 LNYIALGVGFLPPFILTICVLLIIRVLKVEIPESGPRDAQRKALTIVIAMIFLLCF 244
QY 234 LPSVSARLYFLWTPSSACDPSVHCALHITLSFTYWNMSMLDPLVYFSPSPF 285
Db 245 LPYHALRTHLVTDWADSCMDLKHATVITLTAAANSFCNFPFLYFAGENF 296

RESULT 11
P2YR_HUMAN
ID P2YR_HUMAN STANDARD; PRT; 373 AA.
AC P47900;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).
GN P2YR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
```

```
RX MEDLINE-96257237; PubMed-8666290;
RA Leon C., Vial C., Cazenave J.-P., Gachet C.;
RT "Cloning and sequencing of a human cDNA encoding endothelial P2Y1
RT purinoceptor.";
RL Gene 171:295-297(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-96158962; PubMed-8579591;
RA Ayvathan K., Tania W., Harbansjit S., Raghibir A.S., Barnard E.A.,
RA Kunapuli S.P.;
RT "Cloning and chromosomal localization of the human P2Y1
RT purinoceptor.";
RL Biochem. Biophys. Res. Commun. 218:783-788(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE-96205320; PubMed-8630005;
RA Janssens R., Communi D., Pirotton S., Samson M., Parmentier M.,
RA Boeynaems J.M.;
RT "Cloning and tissue distribution of the human P2Y1 receptor.";
RL Biochem. Biophys. Res. Commun. 221:588-593(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Leon C., Vial C., Weber J., Cazenave J.-P., Gacher C.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 95-373 FROM N.A., FUNCTION, AND INHIBITION STUDIES.
RC TISSUE=Platelet;
RX MEDLINE-98113162; PubMed-9442040;
RA Jin J., Daniel J.L., Kunapuli S.P.;
RT "Molecular basis for ADP-induced platelet activation. II. The P2Y1
RT receptor mediates ADP-induced intracellular calcium mobilization and
RT shape change in platelets.";
RL J. Biol. Chem. 273:2030-2034(1998).
CC -1- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
CC ATP AND ADP. IN PLATELETS BINDING TO ADP LEADS TO MOBILIZATION OF
CC INTRACELLULAR CALCIUM IONS VIA ACTIVATION OF PHOSPHOLIPASE C, A
CC CHANGE IN PLATELET SHAPE, AND PROBABLY TO PLATELET AGGREGATION.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- INDUCTION: REPRESSED BY THE P2Y1 RECEPTOR-SPECIFIC ANTAGONISTS
CC A3P5PS, A3P5P AND A2P5P. THESE INHIBIT CALCIUM ION MOBILIZATION
CC AND SHAPE CHANGE IN PLATELETS.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
EMBL; Z49205; CAA89066.1; -.
DR EMBL; U42030; AAA97873.1; -.
DR EMBL; U42029; AAA97872.1; -.
DR EMBL; S81950; AAB47091.1; -.
DR EMBL; AJ006945; CAA07339.1; -.
DR EMBL; AF018284; AAB94556.1; -.
DR HSSP; P34996; 1DDD.
DR GCRDB; GCR_1056; -.
DR GCRDB; GCR_1884; -.
DR GCRDB; GCR_1948; -.
DR GCRDB; GCR_1949; -.
DR MIM; 601167; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_Rhodpsn.
DR PROSITE; PS00237; G_PROTEIN_RECF_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECF_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Platelet;
KW Blood coagulation.
FT DOMAIN 1 52 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 53 74 1 (POTENTIAL).
```



```
P2Y9_HUMAN
ID AC P2Y9_HUMAN STANDARD; PRT; 370 AA.
AC Q99677; O15132;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE P2Y purinoceptor 9 (P2Y9) (Purinergic receptor 9) (G protein-coupled
DE receptor GPR23) (P2Y5-like receptor).
GN GPR23 OR P2RY9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=97225799; PubMed=9073069;
RA O'Dowd B.F., Nguyen T., Jung B.P., Marchese A., Cheng R.,
RA Heng H.H.Q., Kolakowski L.F. Jr., Lynch K.R., George S.R.;
RT "Cloning and chromosomal mapping of four putative novel human
RT G-protein-coupled receptor genes.";
RL Gene 187:75-81(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Bolm S.K., Khitin L.M., Payan D.P., Bunnett N.W.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Janssens R., Boeynaems J.M., Godart M., Communi D.;
RT "Cloning of a human heptahelical receptor closely related to the P2Y5
RT receptor.";
RL Biochem. Biophys. Res. Commun. 236:106-112(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: NOT DETECTED IN THE BRAIN REGIONS THALAMUS,
CC PUTAMEN, CAUDATE, FRONTAL CORTEX, PONS, HYPOTHALAMUS, HIPPOCAMPUS.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U66578; AAC51301.1; -
DR EMBL; U90323; AAB62087.1; -
DR EMBL; U90322; AAB62088.1; -
DR HSSP; AF005419; AAB66322.1; -
DR GCRdb; GCR_1323; -
DR GCRdb; GCR_1324; -
DR GCRdb; GCR_1912; -
DR MIN; 300086; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 43 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 44 64 1 (POTENTIAL).
FT DOMAIN 65 73 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 74 94 2 (POTENTIAL).
FT DOMAIN 95 112 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 113 133 3 (POTENTIAL).
FT DOMAIN 134 155 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 156 176 4 (POTENTIAL).
FT DOMAIN 177 203 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 204 224 5 (POTENTIAL).
FT DOMAIN 225 254 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 255 275 6 (POTENTIAL).
FT DOMAIN 276 294 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 295 315 7 (POTENTIAL).
```

```
FT DOMAIN 316 370 CYTOPLASMIC (POTENTIAL).
FT DISULFID 111 188 BY SIMILARITY.
FT CARBOHYD 15 15 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 24 24 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 192 192 F -> L (IN REF. 3).
SQ SEQUENCE 370 AA; 41895 MW; 20857F52A3929E48 CRC64;

Query Match 18.7%; Score 346.5; DB 1; Length 370;
Best Local Similarity 31.2%; Pred. No. 1.6e-17;
Matches 95; Conservative 52; Mismatches 128; Indels 29; Gaps 11;

QY 24 VAFVLGALGNVAGLCGFCFHMKTWKPSVYLFNLAADFLMIMICLPFRDYLRHRHAF 83
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DQ 48 VVFILGLITNSVSLFVFCFRMKRSETAIFITNLAVDLLFVCTLPFKI-FYFNFRHPF 106
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 84 GDIPCRVGLFTLMMNRAGSIVFLTVAAADRFKVVHPHVHANTSTRVAAGIVCT-LNAL 142
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DQ 107 GDTLKISGTAFLTNIYGSMLFCTICISVDRELAIVYFERS-RTIRTRNSAIVCAGVWIL 165
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 143 VILGTVYLLLENHLCVQETAVSC-ESETMESANGWHDIMFOLE-----FEMPLGIILF 194
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DQ 166 VLSGISASLSFTTNVNNATTTCPEGF--SKRVWKTVLSKITIPFVGVGFIIPILNVS 222
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 195 CSFKIVSLRRRQQLAR-QARMKKATREFIMVVAIVFITCYLPVSVSARLYFLWTVPSSAC- 252
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DQ 223 CSSVVLRLRPATLSQIGTNKKVLMITVHMAVFFVCFVPYNS--VLFYALVRSQAI 280
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 253 -----DPSVGHALHITLSFTWNSMLDPLVYFSSPSPK-FYNKLKICSLKPKQPGHISK 306
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DQ 281 TNCELEFAKIMYPITLCLATLNCDFPFYIYFTLESQKSFYINAH-----RMESLFK 335
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 307 TORP 310
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DQ 336 TETP 339

RESULT 14
P2Y5_CHICK
ID P2Y5_CHICK STANDARD; PRT; 308 AA.
AC P32250;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE P2Y purinoceptor 5 (P2Y5) (Purinergic receptor 5) (6H1).
GN P2RY5.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=T-cell;
RX MEDLINE=93329058; PubMed=8393036;
RA Kaplan M.H., Smith D.I., Sundick R.S.;
RT "Identification of a G protein coupled receptor induced in activated
RT T cells.";
RL J. Immunol. 151:628-636(1993).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: INDUCED IN ACTIVATED T-CELLS.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L06109; AAB06587.1; -
```

```

DR HSP; P34996; 1DDD.
DR GCRDb; GCR_0689; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR01067; P2Y5RNOCPTR.
DR PRINTS; PR01157; P2PURNOCEPTR.
DR PROSITE; PS00237; G_PROTEIN_RECPT_FL1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECPT_FL2_1; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein; Lipoprotein;
KW Palmitate.
FT DOMAIN 1 16 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 17 43 1 (POTENTIAL).
FT DOMAIN 44 52 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 53 76 2 (POTENTIAL).
FT DOMAIN 77 89 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 90 109 3 (POTENTIAL).
FT DOMAIN 110 130 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 131 151 4 (POTENTIAL).
FT DOMAIN 152 178 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 179 206 5 (POTENTIAL).
FT DOMAIN 207 224 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 225 250 6 (POTENTIAL).
FT DOMAIN 251 269 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 270 289 7 (POTENTIAL).
FT DOMAIN 290 308 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 5 5 N-LINKED (GLCNAC. .) (POTENTIAL).
FT LIPID 281 281 PALMITATE (BY SIMILARITY).
FT DISULFID 86 165 POTENTIAL.
SQ SEQUENCE 308 AA; 35597 MW; 4214E96963B6F7D CRC64;

Query Match 18.3%; Score 339; DB 1; Length 308;
Best Local Similarity 30.3%; Pred. No. 4.5e-17;
Matches 91; Conservative 56; Mismatches 123; Indels 30; Gaps 10;

QY 24 VAFVGLGNGVALCGCFHMKTKPVTYVFNLAVALDELLMCLPRTDYLLRRHWAF 83
Db 23 MVEVLGIANCVAIYIFETFLKVRNETTYMLNLAISDLFLVFPRI-YFFVVRNMPF 81

QY 84 GDPICRVGLFTLANNRAGSIVFTLVVAADRYKVVHPHVAVTISTVAAGIVC-TLWAL 142
Db 82 GDVLCKISVILFYTNMNGSILFLTCISVDRELAIVHFRS-KTLRTKNARIVCAVWIT 140

QY 143 VILGTV---YLLLENHLCVQETAVSCSFIMESANGWHIDMFQI-----EFFMPLGII 192
Db 141 VLAGSTPASFPQSNRNQNTQRTCFENF---PESTWKTLSRVIRFIEIVGFPIPLN 197

QY 193 LFCSEKTVMSLRRLRQQLAR-QARKKATRTMVAIVFICYLP-SVSARLYFLW---TV 247
Db 198 VTCSTMVRLTNKPLTLRSRNLKSKKVLKMFVHLVIFCFVFPYNTLILYSLMRQTW 257

QY 248 PSSACDPSVHGALHITLSFTYMNMLDPLVYFSSPFPKPKYFKLKICSLKPKQPGHSHK 307
Db 258 INCSTVAVRTMYPVTLCIAVSNCCFPDIYYFTSDINSE-----LDKKQOVHQT 308

RESULT 15
GPRH_HUMAN STANDARD; PRT; 339 AA.
AC Q13304;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative G protein-coupled receptor GPR17 (R12).
GN GPR17.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96145150; PubMed=8558062;
RA Report C.J., Schweickart V.L., Chantray D., Eddy R.L. Jr., Shows T.B.,
RA Godiska R., Gray P.W.;

```

```

RT "New members of the chemokine receptor gene family.";
RN J. Leukoc. Biol. 59:18-23(1996).
RP SEQUENCE FROM N.A.
RC TISSUE-Hippocampus;
RX MEDLINE=98181695; PubMed=9523551;
RA Blaesiuss R.H., Weber R.G., Lichter P., Ogilvie A.;
RT "A novel orphan G protein-coupled receptor primarily expressed in the
RT brain is localized on human chromosomal band 2q21.";
RL J. Neurochem. 70:1357-1365(1998).
CC -|- FUNCTION: ORPHAN RECEPTOR. POTENTIAL CHEMOKINE RECEPTOR.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein.
CC -|- TISSUE SPECIFICITY: PRIMARILY EXPRESSED IN BRAIN.
CC -|- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; U33447; AAB16746.1; -.
DR EMBL; Z94154; CAB08107.1; -.
DR HSP; P34996; 1DDD.
DR GCRDb; GCR_1968; -.
DR GCRDb; GCR_2597; -.
DR MIM; 603071; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECPT_FL1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECPT_FL2_1; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 36 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 37 57 1 (POTENTIAL).
FT DOMAIN 58 64 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 65 85 2 (POTENTIAL).
FT DOMAIN 86 105 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 106 126 3 (POTENTIAL).
FT DOMAIN 127 147 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 148 168 4 (POTENTIAL).
FT DOMAIN 169 195 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 196 216 5 (POTENTIAL).
FT DOMAIN 217 232 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 233 253 6 (POTENTIAL).
FT DOMAIN 254 280 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 281 301 7 (POTENTIAL).
FT DOMAIN 302 339 CYTOPLASMIC (POTENTIAL).
FT DISULFID 104 181 BY SIMILARITY.
FT CARBOHYD 14 14 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 176 176 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 254 254 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 339 AA; 37826 MW; BD2AEDB2FEBF15E1 CRC64;

Query Match 18.1%; Score 336; DB 1; Length 339;
Best Local Similarity 30.5%; Pred. No. 7.9e-17;
Matches 97; Conservative 54; Mismatches 153; Indels 14; Gaps 9;

QY 7 CRIEGDTISQVMPPLLIIVAFVGLGNGVALCGCFHMKTKPSTVYLFNLAVALDFLLMI 66
Db 23 CGQETPLENMLFASFYLLDFLALVGNLTALWLFIRDHKSCTPANVFLMHLAVALDSLVL 82

QY 67 CLPRTDYLLRRHWAFGDPICRVGLFTLANNRAGSIVFTLVVAADRYKVVHPHVAVT 126
Db 83 VLPRLTYVHFSGNHWPFGEIACRLTGFLYLNMYASIFYLTCISADRELAIVHP---VKS 139
QY 127 ISTR--VAAGIVCT-LWALVILGTVYVLLLENHLCVQETAVSCSFIMESANGWHIDMFQ 183
Db 140 LKLRRLYLAHLACAFLLWVAVAMAPLLVSPQTQVTHVTVCLOLYREKASHHALVSLAV 199
QY 184 EFFMPLGIILFCSEKTVMSLRRLRQQLARQARKM-KATRTMVAIVFITCYLP-SVSARL 241

```

Db 200 AFTFPFITVTCTYLLIIRSL--RQGLRVEKRLKTKAVRMIAIVLAIFLVCFVPYHVNRSV 257
QY 242 YEL-WTVPSACDPSVHGAL--HITLSFTYXNSMLDPLVYFSSPSPFKFYNNKLKICSUK 298
Db 258 YVLHYRSHGASCATORILALANRITSCLTSLNGALDPIMYFFVAEKFRHALCNL-LCGKR 316
QY 299 PKQPGHSGTKORPEEMPIS 316
Db 317 LKGPPPSFEGKTNESSLS 334

Search completed: October 29, 2002, 04:52:57
Job time : 27 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2002, 04:51:26 ; Search time 44 seconds
(without alignments)
755.612 Million cell updates/sec

Title: US-09-886-041-2

Perfect score: 1853

Sequence: 1 MYNGSCRIEGDTISQVMP.....ANSFQSDGQWDPIHVEWH 346

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	880.5	47.5	387	2 I69202	G protein-coupled
2	364	19.6	352	2 S33723	G protein-coupled
3	362.5	19.6	373	2 JC4162	P2Y receptor - bov
4	354	19.1	373	2 JC4737	G protein-coupled
5	343.5	18.5	370	2 JC5549	heptahelical p2y5-
6	339	18.3	308	2 I50241	G protein-coupled
7	324	17.5	344	2 T09508	Intron 17 purinerg
8	320	17.3	373	2 A47556	ATP receptor p2u -
9	310.5	16.8	420	2 I51667	thrombin receptor
10	309.5	16.7	363	2 I57940	somatostatin recep
11	307	16.6	391	2 C41795	somatostatin recep
12	305.5	16.5	391	2 A41795	somatostatin recep
13	305.5	16.5	391	2 A39297	somatostatin recep
14	304.5	16.4	359	2 S15403	angiotensin II rec
15	302.5	16.3	369	2 B41795	somatostatin recep
16	302	16.3	364	2 JQ1488	bradykinin B2 rece
17	302	16.3	399	2 I48705	proteinase activat
18	300.5	16.2	369	2 A45291	somatostatin recep
19	300.5	16.2	369	2 D41795	somatostatin recep
20	300	16.2	388	2 I55450	G protein-coupled
21	300	16.2	384	2 A47249	brain-specific som
22	300	16.2	428	2 A44021	somatostatin recep
23	299.5	16.2	369	2 JC2083	somatostatin recep
24	299	16.1	388	2 JN0605	somatostatin recep
25	298.5	16.1	418	2 A46226	somatostatin recep
26	298.5	16.1	432	2 A43448	thrombin receptor
27	298	16.1	428	2 S30508	probable G protein
28	297	16.0	397	2 S66518	proteinase-activat
29	296.5	16.0	346	2 S29248	somatostatin recep

30	296	16.0	359	2 A48857	angiotensin II rec
31	295	16.0	375	2 A54946	p-2U nucleotide re
32	295	15.9	384	2 JC4629	somatostatin recep
33	294	15.9	372	2 I38532	delta opioid recep
34	292.5	15.8	361	2 JC5653	G protein-coupled
35	292	15.8	371	2 JC5498	angiotensin II rec
36	291.5	15.7	359	2 S44425	angiotensin II rec
37	291.5	15.7	359	2 JC2134	delta opioid recep
38	290	15.7	372	2 S34592	angiotensin II rec
39	289	15.6	359	2 JC1104	chemokine (C-C) re
40	288	15.5	355	2 A45177	delta opioid recep
41	288	15.5	372	2 B48227	G protein-coupled
42	287.5	15.5	365	2 S68208	probable chemotatr
43	287.5	15.5	371	2 JC5796	angiotensin II rec
44	286.5	15.5	359	2 JH0621	allatostatin recep
45	286.5	15.5	423	2 JC7677	

ALIGNMENTS

RESULT 1

I69202

G protein-coupled receptor HM74 - human

C;Species: Homo sapiens (man)

C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 19-May-2000

C;Accession: I69202

R;Nomura, H.; Nielsen, B.W.; Matsushima, K.

Int. Immunol. 5, 1239-1249, 1993

A;Title: Molecular cloning of cDNAs encoding a LD78 receptor and putative leukocyte c

A;Reference number: I54751; MUID:94092629

A;Accession: I69202

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-387 <RES>

A;Cross-references: GB:D10923; NID:g219866; PIDN:BAA01721.1; PID:g219867

C;Genetics:

A;Gene: HM74

C;Superfamily: G protein-coupled receptor 4

Query Match 47.5%; Score 880.5; DB 2; Length 387;
Best Local Similarity 52.2%; Pred. No. 5.1e-74;
Matches 178; Conservative 49; Mismatches 107; Indels 7; Gaps 4;

QY	5	SCCRIEGDTISQVMP	LIIVAFVLGALGV	ALCGCFHMKTKP	STVYLFNLAVADFL	64
Db	17	NCCVFRDDFAKVL	PPVVLGLEIFGL	LGNGLALWPCFH	LKSWKSSRIELFNL	76
QY	65	MICLPFRDYYLRR	HWAFGDI	PCRVGLFTLAMN	RAGSI	124
Db	77	IICLPFVMDYYVR	SDNFGDIP	CLRVLFMFAN	RQGSIIFLIVAVD	136
QY	125	NTISTRVAAGIV	CTLMALVIGTVY	LLLENHLCVQET	AVSCSFIMESANG	184
Db	137	NKISNMTAAIIS	CLLWGITVGLT	VHLKKLLIQNG	PANVCISFSICTH	196
QY	185	FFPLGLILCSK	VIWSLRRQOL	ARQARKATRIM	VAIVFICVLP	244
Db	197	FLUPLGLILCS	ARIWLSLRQ	QMDRHAKIKR	AITIMVAIVFV	255
QY	245	WTPVSSA---CD	--PSVHGALH	ITLSFTYNN	MDPLVYFSSP	299
Db	256	WLLHTSTQNC	EVYRVDLAFF	ITLSFTYNN	MDLPVYFSSP	315
QY	300	KQPGHSKTQ	PEEMPISN	GLRRSCIS	IVANSFQSDGQ	340
Db	316	KMTGPDNNR	STSVELTGD	PNKT-RGA	PEALMANS	355

RESULT 2

S33733

G protein-coupled receptor - chicken

C;Species: Gallus gallus (chicken)

	Best Local Similarity	27.5%;	Pred. No. 6.6e-26;
	Matches	92;	Conservative
	68;	Mismatches	140;
	Indels	35;	Gaps
Qy	18	MPPLLIIVAFVLGALGVLCGFCFHMKTWKPSVYVLEFNLA	VADFLMICLPFRDYYLR 77
Db	54	LPAVYILVFIIGFGLNSVAIWMVFHMKPWSGISVYFNFNLA	DAFLYVTLFALIFYFN 113
Qy	78	RHRWAFGIDPCRYGLETFLAMNRAGSIVFLTVAAADRYFKV	WHPHPHVAVNTISTRVAAGIVC 137
Db	114	KTDWIFGDAMCKLQRFIFHVLNYSILFLTCISAHRYSGV	YVPLKSLGRLLKKNAVYISV 173
Qy	138	TLNALVLGTIVYVLLLENHLCVQET-AVSC-----	SSFTMESANGWHDIMFQLEPFM 187
Db	174	LVLLIVVVGISPIIFYSGTGIRKNKTICTVDTTSDYLSR	SYFIYSM-----CTTVAMFCV 228
Qy	188	PLGTILFCSFKIYVLSRRRQOLARQARMKATFIMVAIV	IFTCVLP-----SVSAR 240
Db	229	PLVLILGCGYILVRLAIYK-DLDNSPLRKSIVLYIVITV	FAVSYPHVMKTNLRLAR 287
Qy	241	LYFLWTVPSSACDPSVHGALHITLSTYWNMSMLDPLV	YVFSSPFPKFYNKIKICSLKPK 300
Db	288	LDF-QTPMCFAFNDRYATVQVTRGLASLNSCDVPLIY	LAGDTFRRLSR----- 337
Qy	301	OPGHSKTRQPEMPIINLGRRCISIVANSFQSQSD	335
Db	338	--ATRKASRRSEANLQSKSDMTNLILSBFKONGD	370
RESULT 4			
JC4737			
N	G	protein-coupled receptor P2Y1 - human	
C	Species:	Homo sapiens (man)	
C	Date:	10-May-1996	#sequence
C	Accession:	JC4737	JC4615; S54253
R	Janssens, R.; Commun, D.; Pirotton, S.; Samson, M.;	Boey	
A	Title:	Cloning and tissue distribution of the human P2Y1 receptor.	
A	Reference number:	JC4737; MUID:96205320	
A	Accession:	JC4737	
A	Molecule type:	DNA	
A	Residues:	1-373 <N>	
A	Cross-references:	GB:S81950; NID:g1839438; PIDN:AAB47091.1; PID:g1839439	
R	Ayvanathan, K.; Webbs, T.E.; Sandhu, A.K.; Athwal, R.S.; Barnard, E.A.;		
Biochem. Biophys. Res. Commun.	218, 783-788, 1996		
A	Title:	Cloning and chromosomal localization of the human P2Y1 purinoceptor	
A	Reference number:	JC4615; MUID:96158962	
A	Accession:	JC4615	
A	Molecule type:	mRNA	
A	Residues:	1-373 <AY>	
A	Cross-references:	GB:U4029; NID:g1147730; PIDN:AAA97872.1; PID:g1147731	
A	Experimental source:	erythro leukemia cells	
R	Leon, C.; Vial, C.; Cazenave, J.; Gachet, C.		
	submitted to the EMBL Data Library, May 1995		
A	Description:	Cloning of a human putative P2Y receptor.	
A	Reference number:	S54253	
A	Accession:	S54253	
A	Status:	preliminary	
A	Molecule type:	mRNA	
A	Residues:	1-137,139-373 <LEO>	
A	Cross-references:	EMBL:249205; NID:g798835; PID:CAA9066.1; PID:g798836	
C	Comment:	This receptor belongs to a family of G protein-coupled receptors.	
C	Genetics:		
A	Gene:	P2Y1; GDB:P2RY1	
A	Cross-references:	GDB:677125; OMIM:601167	
A	Map position:	3pter-3qter	
C	Superfamily:	ATP receptor P2u	
C	Keywords:	G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane	
F	88-111/Domain:	transmembrane #status predicted <TM1>	
F	124-152/Domain:	transmembrane #status predicted <TM3>	
F	171-191/Domain:	transmembrane #status predicted <TM4>	
F	214-237/Domain:	transmembrane #status predicted <TM5>	

THIS PAGE BLANK (USPIC,

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2002, 04:47:16 ; Search time 25 Seconds
(without alignments)
338.050 Million cell updates/sec

Title: US-09-886-041-2
Perfect score: 1853
Sequence: 1 MYNGSCRIEQTISQVMP.....ANSFQSDGWDQPHIVEWH 346

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	529	28.5	423	US-08-955-713-2	Sequence 2, Appli
2	510	27.5	476	US-08-955-713-4	Sequence 4, Appli
3	451.5	24.4	319	US-09-130-749-2	Sequence 2, Appli
4	451.5	24.4	319	US-09-130-749-2	Sequence 2, Appli
5	371	20.0	362	US-08-513-974B-374	Sequence 374, App
6	362.5	19.6	373	US-08-559-524A-4	Sequence 4, Appli
7	362.5	19.6	373	US-08-749-707-4	Sequence 4, Appli
8	346.5	18.7	370	US-08-781-250-2	Sequence 2, Appli
9	336	18.1	339	US-08-153-848-44	Sequence 3, Appli
10	336	18.1	339	US-08-812-871-3	Sequence 4, Appli
11	336	18.1	339	US-09-299-843A-44	Sequence 44, Appl
12	336	18.1	339	US-09-088-337B-44	Sequence 44, Appl
13	336	18.1	339	PCT-US93-11153-44	Sequence 44, Appl
14	336	18.1	339	PCT-US95-07180-2	Sequence 2, Appli
15	328.5	17.7	302	US-08-467-948A-30	Sequence 30, Appl
16	328.5	17.7	302	US-08-467-947A-30	Sequence 30, Appl
17	327	17.6	309	US-09-422-869-20	Sequence 20, Appl
18	325	17.5	344	US-08-467-948A-8	Sequence 8, Appli
19	325	17.5	344	US-08-467-947A-8	Sequence 8, Appli
20	314	16.9	373	US-08-513-974B-373	Sequence 373, App
21	308	16.6	325	US-08-118-270-51	Sequence 51, Appl
22	308	16.6	325	PCT-US93-08528-51	Sequence 51, Appl
23	307	16.6	391	US-07-816-283-4	Sequence 4, Appli
24	307	16.6	391	US-08-417-103-4	Sequence 4, Appli
25	307	16.6	395	US-08-097-938-5	Sequence 5, Appli
26	307	16.6	395	US-08-476-000-5	Sequence 5, Appli
27	307	16.6	395	US-08-472-840-5	Sequence 5, Appli

28	307	16.6	395	2	US-08-476-976-5	Sequence 5, Appli
29	307	16.6	395	3	US-08-474-410-5	Sequence 5, Appli
30	306	16.5	398	1	US-08-097-938-6	Sequence 6, Appli
31	306	16.5	398	1	US-08-476-000-6	Sequence 6, Appli
32	306	16.5	398	1	US-08-472-840-6	Sequence 6, Appli
33	306	16.5	398	2	US-08-476-976-6	Sequence 6, Appli
34	306	16.5	398	3	US-08-474-410-6	Sequence 6, Appli
35	306	16.5	398	4	US-08-486-673B-6	Sequence 6, Appli
36	305.5	16.5	369	4	US-08-120-601B-9	Sequence 9, Appli
37	305.5	16.5	391	1	US-07-816-283-2	Sequence 2, Appli
38	305.5	16.5	391	1	US-08-417-103-2	Sequence 2, Appli
39	305.5	16.5	391	1	US-08-417-103-14	Sequence 14, Appl
40	305	16.5	337	4	US-09-044-404A-2	Sequence 2, Appli
41	303.5	16.4	391	4	US-08-120-601B-8	Sequence 8, Appli
42	302.5	16.3	369	1	US-07-816-283-6	Sequence 6, Appli
43	302.5	16.3	369	1	US-08-417-103-6	Sequence 6, Appli
44	302.5	16.3	369	1	US-08-417-103-16	Sequence 16, Appl
45	302	16.3	395	4	US-08-486-673B-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-955-713-2
; Sequence 2, Application US/08955713
; Patent No. 5955308
; GENERAL INFORMATION:
; APPLICANT: SATHE, GANESH
; APPLICANT: MOONEY, JEFFREY
; APPLICANT: BERGSMAN, DEBK
; APPLICANT: HALSEY, WENDY
; TITLE OF INVENTION: CDNA CLONE HE0AD54 THAT ENCODES A HUMAN 7-TRANS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/955,713
FILING DATE: 23-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/050,124
FILING DATE: 18-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70087
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-955-713-2

Query Match 28.5%; Score 529; DB 2; Length 423;
Best Local Similarity 39.2%; Pred. No. 3.2e-38;
Matches 115; Conservative 48; Mismatches 108; Indels 22; Gaps 7;


```
Db 6 CSAPSTVATAVGVLGLEGGLGNAVALWTFELFRVRWKPYAVYLLNALADLLAA 65
QY 67 CLPFRDYILRRHWAFGDIPCRVGLFTLAMNRAGSIVELTVVAADRYKVVHPHNAVNT 126
Db 66 CLPFLAAYLSLQAHWLGVRGCHALRFLDLDSRVGMAFLAAVALDRYLVRVHPKLVNL 125
QY 127 ISTRVAGIVCTLWALVILGTVYLLLENHLCVQETAVSCSFIMESANG-----WHDIMF 181
Db 126 LSPQAALGVSGVLWLLMVALTCFGLLISE--AAQNSTRCHSF-YSRADGSGFSIIWOEALS 182
QY 182 QLEFFPPLGILFCSPKIWSLRR--QQLARQARMKKATRFIMVVAIVETCYLPSVSAR 240
Db 183 CLQFVLPFGGLIVFCNAGIIRALQKRLREPEKPKLQRAQALVTLVVVLFALCFLPCFLAR 242
QY 241 --LYFLWTVPSSACDPSVHGALHITLSFTYMNMLDPLVYFSSPFPKFNKL 292
Db 243 VLMHIFQNLGSCRALCAVAHSDVTGSLTYLHSLVNPVYVCFSSPTFRSSYRRV 296

RESULT 4
US-09-130-749-2
; Sequence 2, Application US/09130749
; Patent No. 6031344
; GENERAL INFORMATION:
; APPLICANT: SHABON, USMAN
; APPLICANT: ELSHOURBAGY, NABIL
; TITLE OF INVENTION: MOLECULAR CLONING OF A 7TM
; TITLE OF INVENTION: RECEPTOR (GPR31A)
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. Box 980
; CITY: Valley Forge
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/130,749
; FILING DATE: 07-AUG-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GP-70513
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 319 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-130-749-2

Query Match 24.4%; Score 451.5; DB 3; Length 319;
Best Local Similarity 34.4%; Pred. No. 1.1e-31;
Matches 101; Conservative 61; Mismatches 121; Indels 11; Gaps 5;

QY 7 CRIGEDTISQVMPPLLIVAEVLGALNGVACGFCFCHMKTKPSTVYLFNLVADEFLMI 66
Db 6 CSAPSTVATAVGVLGLEGGLGNAVALWTFELFRVRWKPYAVYLLNALADLLAA 65
```

```
QY 67 CLPFRDYILRRHWAFGDIPCRVGLFTLAMNRAGSIVELTVVAADRYKVVHPHNAVNT 126
Db 66 CLPFLAAYLSLQAHWLGVRGCHALRFLDLDSRVGMAFLAAVALDRYLVRVHPKLVNL 125
QY 127 ISTRVAGIVCTLWALVILGTVYLLLENHLCVQETAVSCSFIMESANG-----WHDIMF 181
Db 126 LSPQAALGVSGVLWLLMVALTCFGLLISE--AAQNSTRCHSF-YSRADGSGFSIIWOEALS 182
QY 182 QLEFFPPLGILFCSPKIWSLRR--QQLARQARMKKATRFIMVVAIVETCYLPSVSAR 240
Db 183 CLQFVLPFGGLIVFCNAGIIRALQKRLREPEKPKLQRAQALVTLVVVLFALCFLPCFLAR 242
QY 241 --LYFLWTVPSSACDPSVHGALHITLSFTYMNMLDPLVYFSSPFPKFNKL 292
Db 243 VLMHIFQNLGSCRALCAVAHSDVTGSLTYLHSLVNPVYVCFSSPTFRSSYRRV 296

RESULT 5
US-08-513-974B-374
; Sequence 374, Application US/08513974B
; Patent No. 6114139
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Hosoya, Masaki
; APPLICANT: Fujii, Ryo
; APPLICANT: Ohtaki, Tetsuya
; APPLICANT: Fukusumi, Shoji
; APPLICANT: Ohgi, Kazuhiro
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
; TITLE OF INVENTION: PRODUCTION, AND USE THEREOF
; NUMBER OF SEQUENCES: 380
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/513,974B
; FILING DATE: 14-SEP-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP95/01599
; FILING DATE: 10-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-093989
; FILING DATE: 19-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-057186
; FILING DATE: 16-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-007177
; FILING DATE: 20-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-326611
; FILING DATE: 28-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-270017
; FILING DATE: 02-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-236357
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-236356
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
```

APPLICATION NUMBER: JP 6-189274
FILING DATE: 11-AUG-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 6-189273
FILING DATE: 11-AUG-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 6-189272
FILING DATE: 11-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S.
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 45753
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 374:
SEQUENCE CHARACTERISTICS:
LENGTH: 362 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-513-974B-374

Query Match 20.08; Score 371; DB 3; Length 362;

Best Local Similarity 29.18; Pred. No. 1.2e-24;

Matches 95; Conservative 60; Mismatches 128; Indels 44; Gaps 7;

QY 18 MPPLLIIVAFVLGALNGVAGLCGFCFHMKTWKPSTVYLFNLAADFLLMICLPFRFTDYLR 77
DB 43 LPTVILVFIIFGLGNSVAIMWFVHMRPWSGISVYMFNLAADFLVLTLPALIFYFN 102
QY 78 RRHAFGDIICRGVGLFTLAMNRAGSIIVLTVAAADRFKVVHPHNAVNTISTRVAAGIVC 137
DB 103 KTDWIFGDMCKLQRFIFHVNLYSILFLCISVHRVTGVVHPLKSLGRKKNAVIVSS 162
QY 138 TLWALVILGTVYLLENHNL-----CVOETA-----VCSFTMESANGWHD 178
DB 163 LVWALVAVIAPILFISGIGVRRNKTTICYDTTADYLRSYFVSMCTVFM----- 214
QY 179 IMFQLEFMPGLIIFCSFKIWSLRRRQOLARQARKKATREFIMVAIVFITCYLP--- 235
DB 215 -----TCIPIVILGCVGLIVKALIVK-DLDNSPLRRKSIYLVIIIVTVFAVSILPFHV 267
QY 236 ----SVSARLYFLWTPSSACDPSVHGALHITLSFTYMNMSMLDPLVYFSSPFPKYNK 291
DB 268 MKTLNLRARLDF-QTPQMCFAFNDRVYATYQVTRGLASLNSCVDPILYFLAGDTFRRLSR 326
QY 292 LKICSLKPKQPGHKTQRPPEMPISNL 318
DB 327 ATRKSRSRSEP--NVQSKSEMTNLIL 351

RESULT 6

US-08-559-524A-4

Sequence 4, Application US/08559524A

Patent No. 5871963

GENERAL INFORMATION:

APPLICANT: Conley, Pamela B.

APPLICANT: Jantzen, Hans-Michael

TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP

STREET: 1800 M Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20036-5869

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/559,524A
FILING DATE: 15-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 044481-5010-00-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-467-7000
TELEFAX: 202-467-7176
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-559-524A-4

Query Match 19.68; Score 362.5; DB 2; Length 373;

Best Local Similarity 27.58; Pred. No. 6.5e-24;

Matches 92; Conservative 68; Mismatches 140; Indels 35; Gaps 7;

QY 18 MPPLLIIVAFVLGALNGVAGLCGFCFHMKTWKPSTVYLFNLAADFLLMICLPFRFTDYLR 77
DB 54 LPVAVILVFIIFGLGNSVAIMWFVHMRPWSGISVYMFNLAADFLVLTLPALIFYFN 113
QY 78 RRHAFGDIICRGVGLFTLAMNRAGSIIVLTVAAADRFKVVHPHNAVNTISTRVAAGIVC 137
DB 114 KTDWIFGDMCKLQRFIFHVNLYSILFLCISAHRYSGVYVPLKSLGRKKNAVIVS 173
QY 138 TLWALVILGTVYLLENHNLCOET-AVSC-----ESFIMESANGWHDIMFQLEFFM 187
DB 174 LVWILVVGISPIIFYSGTGIRKNKTTICYDTTSDYLRSYFISM-----CTTVAMFCV 228
QY 188 PLGLIFCSFKIWSLRRRQOLARQARKKATREFIMVAIVFITCYLP-----SVSAR 240
DB 229 PLVILGCVGLIVKALIVK-DLDNSPLRRKSIYLVIIIVTVFAVSIPFHVMTMNLRR 287
QY 241 LYFLWTPSSACDPSVHGALHITLSFTYMNMSMLDPLVYFSSPFPKYNKLCISLKP 300
DB 288 LDF-QTPQMCFAFNDRVYATYQVTRGLASLNSCVDPILYFLAGDTFRRLSR----- 337
QY 301 QPGHKTQRPPEMPISNLRRRSCISVANSFOSQD 335
DB 338 --ATRKASRRSEANLQSKSEDMTNLILSEFKQNGD 370

RESULT 7

US-08-749-707-4

Sequence 4, Application US/08749707

Patent No. 6063582

GENERAL INFORMATION:

APPLICANT: Conley, Pamela B.

APPLICANT: Jantzen, Hans-Michael

TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP

STREET: 1800 M Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20036-5869

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/749,707

ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,843A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA: US 09/088,337
FILING DATE: 01-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/153,848
FILING DATE: 17-NOV-1993
PRIOR APPLICATION DATA: US 07/977,452
FILING DATE: 17-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Jill E. Uhl
REGISTRATION NUMBER: 43,213
REFERENCE/DOCKET NUMBER: 27866/32059B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX:
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 339 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-299-843A-44

Query Match 18.1%; Score 336; DB 3; Length 339;
Best Local Similarity 30.5%; Pred. No. 1.1e-21;
Matches 97; Conservative 54; Mismatches 153; Indels 14; Gaps 9;

QY 7 CRTEGDTISOVMPDLLIVAFVLGALNGVALCGCFHMKTKPSTVYLFNLAVADFLMI 66
Db 23 CGQETPLENMLFASFYLLDILALVGNLTALWLFIRDHKSGTTPANVFLMHLAVADLSVCL 82
QY 67 CLPFRDYYLRRRHAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVWHPHVAVT 126
Db 83 VLPTRLVYHFSGNHWPGEIACRLTGLFYLNMVASYIFLTCISADRELAIVHP---VKS 139
QY 127 ISTR--VAAGIVCT-LMALVILGTVYLLLENHLCVQETAVSCSFIMESANGWHDIMFOL 183
Db 140 LKURRPLYAHLACAFLLWVAVAMAPLLVSPQTVQTNHTVVCLQLYREKASHHALVSLAV 199
QY 184 EFEMPLGIILFCSEFKIVMSLRRROOLARQARMK-KATREFIMVVAIVFITCYLP-SVSARL 241
Db 200 AFTFPFTTTCYLLIIRSL--ROGLRVERLKTAVRMIAIVLAIFLVCVPVHVNRSV 257
QY 242 YFL-WTVPSACDPSVHGAL--HITLSFTYMNMSMLDPLVYFFSSPPKFKYNKLGICSLK 298
Db 258 VYLHYRSHGASCATORILALANRITSLTSLNGALDPIIMYFFVAEKRHALCNL-LGCKR 316
QY 299 PKQPGHKSITORPEEMPIS 316
Db 317 LKGPSPSFEKGTNESSLS 334

RESULT 12
US-09-088-337B-44
; Sequence 44, Application US/09088337B
; Patent No. 6348574
; GENERAL INFORMATION:
; APPLICANT: Godiska, Ronald
; Gray, Patrick W.
; Schwellkart, Vicki L.
; TITLE OF INVENTION: No. 6348574el Seven Transmembrane Receptors

NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/088,337B
FILING DATE: 01-JUN-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/153,848
FILING DATE: 17-NOV-1993
APPLICATION NUMBER: US 07/977,452
FILING DATE: 17-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 6348574and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31794
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 339 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-09-088-337B-44

Query Match 18.1%; Score 336; DB 4; Length 339;
Best Local Similarity 30.5%; Pred. No. 1.1e-21;
Matches 97; Conservative 54; Mismatches 153; Indels 14; Gaps 9;
QY 7 CRTEGDTISOVMPDLLIVAFVLGALNGVALCGCFHMKTKPSTVYLFNLAVADFLMI 66
Db 23 CGQETPLENMLFASFYLLDILALVGNLTALWLFIRDHKSGTTPANVFLMHLAVADLSVCL 82
QY 67 CLPFRDYYLRRRHAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVWHPHVAVT 126
Db 83 VLPTRLVYHFSGNHWPGEIACRLTGLFYLNMVASYIFLTCISADRELAIVHP---VKS 139
QY 127 ISTR--VAAGIVCT-LMALVILGTVYLLLENHLCVQETAVSCSFIMESANGWHDIMFOL 183
Db 140 LKURRPLYAHLACAFLLWVAVAMAPLLVSPQTVQTNHTVVCLQLYREKASHHALVSLAV 199
QY 184 EFEMPLGIILFCSEFKIVMSLRRROOLARQARMK-KATREFIMVVAIVFITCYLP-SVSARL 241
Db 200 AFTFPFTTTCYLLIIRSL--ROGLRVERLKTAVRMIAIVLAIFLVCVPVHVNRSV 257
QY 242 YFL-WTVPSACDPSVHGAL--HITLSFTYMNMSMLDPLVYFFSSPPKFKYNKLGICSLK 298
Db 258 VYLHYRSHGASCATORILALANRITSLTSLNGALDPIIMYFFVAEKRHALCNL-LGCKR 316
QY 299 PKQPGHKSITORPEEMPIS 316
Db 317 LKGPSPSFEKGTNESSLS 334

RESULT 13
PCT-US93-11153-44
; Sequence 44, Application PC/TUS9311153
; GENERAL INFORMATION:

Query Match	17.7%	Score	328.5;	DB 2;	Length	302;					
Best Local Similarity	30.0%;	Pred. No.	4.4e-21;								
Matches	86;	Conservative	58;	Mismatches	122;	Indels	21;	Gaps	9;		
QY	24	VAFVLGALGNV	ALCGCFHMKP	STVFLNF	LNLADELL	MICLPRTDY	LLRRHWAF	83			
Db	21	MVFVLGIANC	VAIYIFFTL	KVRENTT	YMLMAIS	DLFLVFT	LPERI-Y	YFVVRNMPF	79		
QY	84	GDPCRVGLT	FLAMNRAGS	IVFLVAA	DRYKVV	PHHVA	NTISTR	VAAAGIVC-TL	142		
Db	80	GDVLCKSV	ILFYTNMGS	ILFLFCIS	VDRFLA	IVHPFRS-K	TLRTRK	NRARIVCV	138		
QY	143	VILGTV---	YLLLENH	LHCVOE	TAVSCSF	IMESANG	WHDIMFQ	-----E	FFEMPLGII	192	
Db	139	VLAGSTPAS	FQSTNRQNN	TEQRCF	ENF---P	STWKT	YLSRIV	IFIEIVG	FFIPLIN	195	
QY	193	LFCSFKI	WLSLRRRQ	QIAR-Q	ARMKKAT	RFIMV	VAIVF	ITCYLP-S	VYSARIFLW---	TV	247
Db	196	VTCSTMV	LRLTLNK	PLTILSR	NKLSKKK	LMFIV	HLVIF	CFCCFV	PNITILY	LSLMRTQW	255
QY	248	PSSACD	PSVHGAL	HITLS	FTYMN	SLDPL	YVYF	SPSP	FPKPKY	NKLKI	294
Db	256	INCSWV	AVATMY	DPVTLC	IAVSN	CCFOP	IVYIFT	SDTN	SELDK	KQOV	302

Search completed: October 29, 2002, 04:55:15
Job time : 28 secs

THIS PAGE BLANK (USP 10)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2002, 01:10:36 ; Search time 1921 Seconds

(without alignments)
11340.200 Million cell updates/sec

Title: US-09-886-041-1

Perfect score: 1041

Sequence: 1 atgtacaacgggtgctg.....acattgttgaggcactga 1041

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_higo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	1041	100.0	1041	6	AX147834	Sequence
2	1041	100.0	1041	6	AX148182	Sequence
3	1041	100.0	1041	6	AX299705	Sequence
4	1041	100.0	1041	9	AF411110	Homo sapi
5	1041	100.0	1050	6	AX338371	Sequence
6	1041	100.0	1050	6	AX338373	Sequence
7	1041	100.0	1083	6	AX305131	Sequence
8	1041	100.0	1730	6	AX277635	Sequence
9	1041	100.0	2331	6	AX299707	Sequence
10	1041	100.0	179172	2	AC026331	Homo sapi
11	1041	100.0	204062	2	AC026333	Homo sapi
12	1037.8	99.7	1104	6	AX338374	Sequence
13	888	85.3	888	6	AX147766	Sequence
14	369.2	35.5	1092	6	AX148194	Sequence
15	367.6	35.3	2051	6	AX335595	Sequence
16	367.6	35.3	2051	9	HUMH74	Human mRNA
17	356.2	34.2	1956	10	MMU300198	Mus muscu
18	356.2	34.2	7100	10	MMU300199	Mus muscu
19	229.6	22.1	545	6	AR098624	Sequence
20	229.6	22.1	545	6	I65543	Sequence
21	188.8	18.1	1272	6	AX148166	Sequence
22	188.8	18.1	1272	6	AX338078	Sequence
23	188.8	18.1	1449	6	AX338076	Sequence
24	188.8	18.1	1594	6	AR075071	Sequence
25	188.8	18.1	1594	6	E31720	CDNA clone
26	188.8	18.1	2429	9	AC098824	Homo sapi
27	188.8	18.1	157758	9	AC098824	Homo sapi
28	188	18.1	1435	6	AR075072	Sequence
29	188	18.1	1435	6	E31721	CDNA clone
30	184.8	17.8	8042	6	AX345829	Sequence
31	181.2	17.4	8042	6	AX345828	Sequence
32	139.2	13.4	101230	9	HSBA517H2	Human DNA
33	139.2	13.4	205272	2	AL353591	Homo sapi
34	136	13.1	560	6	AX224752	Sequence
35	136	13.1	2061	6	AX334111	Sequence
36	136	13.1	2061	6	AX335616	Sequence
37	136	13.1	2061	9	HSU65402	Human seven
38	130.6	12.5	29689	10	AF140708	Mus muscu
39	121.2	11.6	30819	10	AF140707	Mus muscu
40	109	10.5	1741	9	HS294154	H.sapiens m
41	109	10.5	1900	9	HSU33447	Human putat
42	109	10.5	1901	6	AR107258	Sequence
43	109	10.5	2070	9	HSP2YL6	H.sapiens m
44	109	10.5	204140	9	AC010976	Homo sapi
45	105	10.1	2183	9	HS294155	H.sapiens m

ALIGNMENTS

RESULT 1
AX147834

LOCUS

AX147834 Sequence 79 from Patent WO0136473.

DEFINITION

AX147834

ACCESSION

AX147834.1

VERSION

GI:14346838

KEYWORDS

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1041)

Vogeli,G., Wood,L.S., Parodi,L.A., Hiebsch,R.R., Lind,P.,

Slightom,J., Schellin,K.A., Kayes,P.S., Bannigan,C.M., Ruff,V.,

Sejltiz,T. and Huff,R.M.

Novel g protein-coupled receptors

Patent: WO 0136473-A 79 25-MAY-2001;

PHARMACIA & UPJOHN COMPANY (US)

FEATURES

Location/Qualifiers

1..1041

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 208 a 294 c 278 g 261 t

ORIGIN

Query Match 100.0%; Score 1041; DB 6; Length 1041;
Best Local Similarity 100.0%; Pred. No. 3.5e-249;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTACAAAGGGTCGTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATGCCCGC 60
|||||
Db 1 ATGTACAAAGGGTCGTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATGCCCGC 60

QY 61 CTGCTCATTTGGCCCTTTGCTGGGGCCTAGGCAATGGGTGCGCCCTGTGGTTTC 120
|||||
Db 61 CTGCTCATTTGGCCCTTTGCTGGGGCCTAGGCAATGGGTGCGCCCTGTGGTTTC 120

QY 121 TGCTTCCACATGAAGACCTCGAAGCCAGCACCTGTTTACCTTTTCAATTTGGCGTGGCT 180
|||||
Db 121 TGCTTCCACATGAAGACCTCGAAGCCAGCACCTGTTTACCTTTTCAATTTGGCGTGGCT 180

QY 181 GATTTCTCTTATGATCTGCCTGGCTTTTTCGGACAGACTATTACCTTCAGACGTAGACAC 240
|||||
Db 181 GATTTCTCTTATGATCTGCCTGGCTTTTTCGGACAGACTATTACCTTCAGACGTAGACAC 240

QY 241 TGGGCTTTTGGGACATTCCTGCCGAGTGGGCTCTTCAGCTTGGCCATGAACAGGGCC 300
|||||
Db 241 TGGGCTTTTGGGACATTCCTGCCGAGTGGGCTCTTCAGCTTGGCCATGAACAGGGCC 300

QY 301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTCGCGACAGGTATTTCAAAGTGGTCCACCCC 360
|||||
Db 301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTCGCGACAGGTATTTCAAAGTGGTCCACCCC 360

QY 361 CACCACCGGTGAACATATCTCCACCCGGTGGCGCTGGCATCTGCTGACCCCTGTGG 420
|||||
Db 361 CACCACCGGTGAACATATCTCCACCCGGTGGCGCTGGCATCTGCTGACCCCTGTGG 420

QY 421 GCCCTGTCTATCTCGGAACAGTGTATCTTTTGTGCGAGAACCATCTCTCGGTGCAAGAG 480
|||||
Db 421 GCCCTGTCTATCTCGGAACAGTGTATCTTTTGTGCGAGAACCATCTCTCGGTGCAAGAG 480

QY 481 ACGGCCCTCTCCTGTGAGAGCTTTTCATCATGGAGTTCGCCCAATGGCTGGCATGACATCATG 540
|||||
Db 481 ACGGCCCTCTCCTGTGAGAGCTTTTCATCATGGAGTTCGCCCAATGGCTGGCATGACATCATG 540

QY 541 TTCCAGCTGGAGTCTTTTATGCCCTCGGCATCATCTTATTTTGCCTCTTCAAGATTGTT 600
|||||
Db 541 TTCCAGCTGGAGTCTTTTATGCCCTCGGCATCATCTTATTTTGCCTCTTCAAGATTGTT 600

QY 601 TGGAGCTGAGGGGAGGACAGCTGSCCAGACAGCTCGGATGAAGAGGCGACCCCG 660
|||||
Db 601 TGGAGCTGAGGGGAGGACAGCTGSCCAGACAGCTCGGATGAAGAGGCGACCCCG 660

QY 661 TTCATCATGTGTGGCAATTGTTCATCATGCTGCTGACCTGCCAGCGTGTCTGTAG 720
|||||
Db 661 TTCATCATGTGTGGCAATTGTTCATCATGCTGCTGACCTGCCAGCGTGTCTGTAG 720

QY 721 CTCATTTCTCTGGAGCGGTGCCCTCGAGTGCCTGGCATCCCTCTGTCCATGGGGCCCTG 780
|||||
Db 721 CTCATTTCTCTGGAGCGGTGCCCTCGAGTGCCTGGCATCCCTCTGTCCATGGGGCCCTG 780

QY 781 CACATAACCTTCAGCTTACCTACATGAACAGCATGCTGGATCCCTGGTGTATTATTT 840
|||||
Db 781 CACATAACCTTCAGCTTACCTACATGAACAGCATGCTGGATCCCTGGTGTATTATTT 840

QY 841 TCAAGCCCTCTCTTCCCAATTTCTACAACAAGCTCAAAATCTGCAGTCTGAAACCCCAAG 900
|||||
Db 841 TCAAGCCCTCTCTTCCCAATTTCTACAACAAGCTCAAAATCTGCAGTCTGAAACCCCAAG 900

QY 901 CAGCCAGGACACTCAAAAACACAAAGCCGGAAGAGATGCCAAATTCGAACCTCGGTGCG 960
|||||
Db 901 CAGCCAGGACACTCAAAAACACAAAGCCGGAAGAGATGCCAAATTCGAACCTCGGTGCG 960

QY 961 AGGAGTTGCATCAGTGGGCAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020
|||||
Db 961 AGGAGTTGCATCAGTGGGCAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020

QY 1021 CACATTTGTTGAGTGGCACTGA 1041
|||||
Db 1021 CACATTTGTTGAGTGGCACTGA 1041

RESULT 2
AX148182
LOCUS AX148182 1041 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 23 from Patent WO0136471.
ACCESSION AX148182
VERSION AX148182.1 GI:14347084
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1041)
AUTHORS Chen,R., Dang,H.T. and Lowitz,K.P.
TITLE Endogenous and non-endogenous versions of human g protein-coupled
receptors
JOURNAL Patent: WO 0136471-A 23 25-MAY-2001;
Arena Pharmaceuticals, Inc. (US)
FEATURES
source
1. .1041
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 208 a 294 c 278 g 261 t
ORIGIN

Query Match 100.0%; Score 1041; DB 6; Length 1041;
Best Local Similarity 100.0%; Pred. No. 3.5e-249;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTACAAAGGGTCGTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATGCCCGC 60
|||||
Db 1 ATGTACAAAGGGTCGTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATGCCCGC 60

QY 61 CTGCTCATTTGGCCCTTTGCTGGGGCCTAGGCAATGGGTGCGCCCTGTGGTTTC 120
|||||
Db 61 CTGCTCATTTGGCCCTTTGCTGGGGCCTAGGCAATGGGTGCGCCCTGTGGTTTC 120

QY 121 TGCTTCCACATGAAGACCTCGAAGCCAGCACCTGTTTACCTTTTCAATTTGGCGTGGCT 180
|||||
Db 121 TGCTTCCACATGAAGACCTCGAAGCCAGCACCTGTTTACCTTTTCAATTTGGCGTGGCT 180

QY 181 GATTTCTCTTATGATCTGCCTGGCTTTTTCGGACAGACTATTACCTTCAGACGTAGACAC 240
|||||
Db 181 GATTTCTCTTATGATCTGCCTGGCTTTTTCGGACAGACTATTACCTTCAGACGTAGACAC 240

QY 241 TGGGCTTTTGGGACATTCCTGCCGAGTGGGCTCTTCAGCTTGGCCATGAACAGGGCC 300
|||||
Db 241 TGGGCTTTTGGGACATTCCTGCCGAGTGGGCTCTTCAGCTTGGCCATGAACAGGGCC 300

QY 301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTCGCGACAGGTATTTCAAAGTGGTCCACCCC 360
|||||
Db 301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTCGCGACAGGTATTTCAAAGTGGTCCACCCC 360

QY 361 CACCACCGGTGAACATATCTCCACCCGGTGGCGCTGGCATCTGCTGACCCCTGTGG 420
|||||
Db 361 CACCACCGGTGAACATATCTCCACCCGGTGGCGCTGGCATCTGCTGACCCCTGTGG 420

QY 421 GCCCTGTCTATCTCGGAACAGTGTATCTTTTGTGCGAGAACCATCTCTCGGTGCAAGAG 480
|||||
Db 421 GCCCTGTCTATCTCGGAACAGTGTATCTTTTGTGCGAGAACCATCTCTCGGTGCAAGAG 480

QY 481 ACGGCCCTCTCCTGTGAGAGCTTTTCATCATGGAGTTCGCCCAATGGCTGGCATGACATCATG 540
|||||
Db 481 ACGGCCCTCTCCTGTGAGAGCTTTTCATCATGGAGTTCGCCCAATGGCTGGCATGACATCATG 540

QY 541 TTCAGCTGGAGTCTTTTATGCCCTCGGCATCATCTTATTTTGCCTCTTCAAGATTGTT 600
|||||
Db 541 TTCAGCTGGAGTCTTTTATGCCCTCGGCATCATCTTATTTTGCCTCTTCAAGATTGTT 600

QY 601 TGGAGCTGAGGGGAGGACAGCTGSCCAGACAGCTCGGATGAAGAGGCGACCCCG 660
|||||
Db 601 TGGAGCTGAGGGGAGGACAGCTGSCCAGACAGCTCGGATGAAGAGGCGACCCCG 660

QY 661 TTCATCATGTGTGGCAATTGTTCATCATGCTGCTGACCTGCCAGCGTGTCTGTAG 720
|||||
Db 661 TTCATCATGTGTGGCAATTGTTCATCATGCTGCTGACCTGCCAGCGTGTCTGTAG 720

QY 721 CTCATTTCTCTGGAGCGGTGCCCTCGAGTGCCTGGCATCCCTCTGTCCATGGGGCCCTG 780
|||||
Db 721 CTCATTTCTCTGGAGCGGTGCCCTCGAGTGCCTGGCATCCCTCTGTCCATGGGGCCCTG 780

QY 781 CACATAACCTTCAGCTTACCTACATGAACAGCATGCTGGATCCCTGGTGTATTATTT 840
|||||
Db 781 CACATAACCTTCAGCTTACCTACATGAACAGCATGCTGGATCCCTGGTGTATTATTT 840

QY 841 TCAAGCCCTCTCTTCCCAATTTCTACAACAAGCTCAAAATCTGCAGTCTGAAACCCCAAG 900
|||||
Db 841 TCAAGCCCTCTCTTCCCAATTTCTACAACAAGCTCAAAATCTGCAGTCTGAAACCCCAAG 900

QY 901 CAGCCAGGACACTCAAAAACACAAAGCCGGAAGAGATGCCAAATTCGAACCTCGGTGCG 960
|||||
Db 901 CAGCCAGGACACTCAAAAACACAAAGCCGGAAGAGATGCCAAATTCGAACCTCGGTGCG 960

QY 961 AGGAGTTGCATCAGTGGGCAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020
|||||
Db 961 AGGAGTTGCATCAGTGGGCAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020

REFERENCE 1 (bases 1 to 1041)
AUTHORS Lee,D.K., Nguyen,T., Lynch,K.R., Cheng,R., Vanti,W.B., Arkhitko,O.,
Lewis,T., Evans,J.F., George,S.R. and O'Dowd,B.F.
TITLE Discovery and mapping of ten novel G protein-coupled receptor genes
JOURNAL Gene 275 (1), 83-91 (2001)
MEDLINE 21458557
PUBMED 11574155
REFERENCE 2 (bases 1 to 1041)
AUTHORS Lee,D.K., Nguyen,T., Lynch,K.R., Cheng,R., Vanti,W.B., Arkhitko,O.,
Lewis,T., Evans,J.F., George,S.R. and O'Dowd,B.F.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-2001) Department of Pharmacology, University of
Toronto, 8 Taddle Creek Rd. Rm. 4353, Toronto, Ont M5S 1A8, Canada
FEATURES
source
1..1041
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/map="12q"
1..>1041
/gene="GPR81"
/product="G protein-coupled receptor"
1..>1041
/gene="GPR81"
1..1041
/gene="GPR81"
/codon_start=1
/product="G protein-coupled receptor"
/protein_id="AAU26481.1"
/db_xref="GI:16566326"
translation="MYNGSCRCIEGDTISQVMPPLLIIVFLGALNGVALCGFCFHM
KWKPTVYLFNLAVDLIMICLPFTDYLLRRHWFAGDIPCRVGLFTLMNRAGS
IVPLTVVAADRYFKVHPHNAVNTISTRVAAGIVCTLWALVILGTWLLLENHLCVQE
TAVSCSFTHESANGHNDINFOLEFPNPIGLIFCSFKIWSLRRLRRQQLARQRMKA
TRIMVAIVFICILYPSVARSARYFLWTPSSACDPSVHGALHITISFTYMNMLDPL
VYFSPSPKPFKNKLIKSLKPKQGHKSTORPEMPISNLGRRSCISVANSFQSQS
DGQWDPHIVEWH"
BASE COUNT 208 a 294 c 278 g 261 t
ORIGIN
Query Match 100.0%; Score 1041; DB 9; Length 1041;
Best Local Similarity 100.0%; Pred. No. 3.5e-249;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTACACGGGTCGTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATCCCGCG 60
DB 1 ATGTACACGGGTCGTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATCCCGCG 60
QY 61 CTGCTCATTTGGCCTTTGTGCTGGCGCACACTAGGCAATGGGTCGCCCTGTGTGTTTC 120
DB 61 CTGCTCATTTGGCCTTTGTGCTGGCGCACACTAGGCAATGGGTCGCCCTGTGTGTTTC 120
QY 121 TGCCTTCACATGAAGACCTGGAAGCCAGCACACTGTTTACCTTTTCAATTTGGCGGTGGCT 180
DB 121 TGCCTTCACATGAAGACCTGGAAGCCAGCACACTGTTTACCTTTTCAATTTGGCGGTGGCT 180
QY 181 GATTTCCTCTTATGATCTGCTGCTTTCGGACACACTATTACCTTCAGACGTAGACAC 240
DB 181 GATTTCCTCTTATGATCTGCTGCTTTCGGACACACTATTACCTTCAGACGTAGACAC 240
QY 241 TGGCCTTTTGGGGACATTCCCTGCGAGTGGGCTCTTCAGTTGGCCATGAACAGGGCC 300
DB 241 TGGCCTTTTGGGGACATTCCCTGCGAGTGGGCTCTTCAGTTGGCCATGAACAGGGCC 300
QY 301 GGGAGCATCGTGTTCCTTAGGTGGTGGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC 360
DB 301 GGGAGCATCGTGTTCCTTAGGTGGTGGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC 360
QY 361 CACCACCGGTGAACTATCTCCACCCGGGTGGCGCTGCATCGTCCACCTCTGG 420
DB 361 CACCACCGGTGAACTATCTCCACCCGGGTGGCGCTGCATCGTCCACCTCTGG 420
QY 421 GCCCTGTCATCTCGGGAACAGTGATATCTTTTGTGAGAACCATCTCTCGTGCAGAGAG 480

Db 421 GCCCTGTCATCTCTGGGAACAGTGATATCTTTTGTGAGAACCATCTCTCGTGCAGAGAG 480
QY 481 ACGGCCCTCTCTCTGTGAGAGCTTCATCATGAGTCGCGCAATGCTGCGATGACATG 540
DB 481 ACGGCCCTCTCTCTGTGAGAGCTTCATCATGAGTCGCGCAATGCTGCGATGACATG 540
QY 541 TTCCAGCTGGAGTTCTTTATGCCCTCGGCGATCATCTTATTTTCTCTCTCAAGATTGTT 600
DB 541 TTCCAGCTGGAGTTCTTTATGCCCTCGGCGATCATCTTATTTTCTCTCTCAAGATTGTT 600
QY 601 TGGAGCTGAGGCGGAGGAGCAGCTGCCAGACAGCGCTGGATGAAGAGGCGACCCGG 660
DB 601 TGGAGCTGAGGCGGAGGAGCAGCTGCCAGACAGCGCTGGATGAAGAGGCGACCCGG 660
QY 661 TTTCATCATGTTGGTGGCAATTGTGTTTCATCATGCTACCTGCGGCGATGCTGCTAG 720
DB 661 TTTCATCATGTTGGTGGCAATTGTGTTTCATCATGCTACCTGCGGCGATGCTGCTAG 720
QY 721 CTCTATTTCTCTGAGCGGTGCCCTCGAGTCGCTGCGATCCCTGTCCATGGGCGCTG 780
DB 721 CTCTATTTCTCTGAGCGGTGCCCTCGAGTCGCTGCGATCCCTGTCCATGGGCGCTG 780
QY 781 CACATAACCTTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTGGTGTATTTT 840
DB 781 CACATAACCTTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTGGTGTATTTT 840
QY 841 TCAAGCCCTCTCTTCCCAATTTCTACAACAGCTCAAAATCTCAGTCTGAAACCCAAAG 900
DB 841 TCAAGCCCTCTCTTCCCAATTTCTACAACAGCTCAAAATCTCAGTCTGAAACCCAAAG 900
QY 901 CAGCAGGACACTCAAAAACAAAGCGGAGAGATGCCAATTTGGAACCTCGGTGCG 960
DB 901 CAGCAGGACACTCAAAAACAAAGCGGAGAGATGCCAATTTGGAACCTCGGTGCG 960
QY 961 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCAGTCTGATGGCAATGGATGCC 1020
DB 961 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCAGTCTGATGGCAATGGATGCC 1020
QY 1021 CACATTGTTGAGTGGCACTGA 1041
DB 1021 CACATTGTTGAGTGGCACTGA 1041
RESULT 5
AX338371 AX338371 1050 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 1 from Patent WO0174904.
DEFINITION AX338371
ACCESSION AX338371
VERSION AX338371.1 GI:18128869
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Majumder,K., Vernet,C.A., Casman,S.J., Wolenc,A.R., Spaderna,S.K.,
Padigaru,M., Mishnu,V.S., Tchernev,V.T., Spytek,K.A., Li,L.,
Baumgartner,J.C. and Gusev,V.Y.
TITLE Novel proteins and nucleic acids encoding same
JOURNAL Patent: WO 0174904-A 1 11-OCT-2001;
Curagen Corporation (US)
FEATURES
source
1..1050
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 211 a 299 c 279 g 261 t
ORIGIN
Query Match 100.0%; Score 1041; DB 6; Length 1050;
Best Local Similarity 100.0%; Pred. No. 3.5e-249;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGTACAACGGGTGCTGCTCCGCATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 60
Db 6 ATGTACAACGGGTGCTGCTCCGCATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 65
QY 61 CTGCTCATTTGGCCCTTTGCTGGGGCACTAGGCAATGGGGTCGCCCTGTGTGTTTC 120
Db 66 CTGCTCATTTGGCCCTTTGCTGGGGCACTAGGCAATGGGGTCGCCCTGTGTGTTTC 125
QY 121 TGCTTCCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTTCAATTTGGCCGTGGCT 180
Db 126 TGCTTCCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTTCAATTTGGCCGTGGCT 185
QY 181 GATTTCCTCTTATGATCTGCCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
Db 186 GATTTCCTCTTATGATCTGCCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 245
QY 241 TGGGCTTTTGGGACATTTCCCTCCGAGTGGGCTCTTACAGTTGGCCATGAACAGGCC 300
Db 246 TGGGCTTTTGGGACATTTCCCTCCGAGTGGGCTCTTACAGTTGGCCATGAACAGGCC 305
QY 301 GGGAGCATCTGTTCCCTTACGGTGGTGGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC 360
Db 306 GGGAGCATCTGTTCCCTTACGGTGGTGGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC 365
QY 361 CACACGCGGTGAACACTATCTCCACCCGGTGGCGCTGGCATGCTCTCCGTCGCAAGAG 420
Db 366 CACACGCGGTGAACACTATCTCCACCCGGTGGCGCTGGCATGCTCTCCGTCGCAAGAG 425
QY 421 GCCCTGTCATCTGGGAAAGTGTATCTTTTGTCTGGAGAACCATCTCTCGTGCAGAG 480
Db 426 GCCCTGTCATCTGGGAAAGTGTATCTTTTGTCTGGAGAACCATCTCTCGTGCAGAG 485
QY 481 ACGGCCCTCTCTGTGAGAGCTTCATCATGGAGTGGGCCAATGGCTGGCATGACATCATG 540
Db 486 ACGGCCCTCTCTGTGAGAGCTTCATCATGGAGTGGGCCAATGGCTGGCATGACATCATG 545
QY 541 TTCAGCTGAGTCTTTATGCCCTCGGCATCATCTTATTTTGTCTCTTCAAGATTGTT 600
Db 546 TTCAGCTGAGTCTTTATGCCCTCGGCATCATCTTATTTTGTCTCTTCAAGATTGTT 605
QY 601 TGGAGCTGAGGGGAGGAGAGCTGGCCAGACAGGCTGGATGAAGAGGGGACCCCG 660
Db 606 TGGAGCTGAGGGGAGGAGAGCTGGCCAGACAGGCTGGATGAAGAGGGGACCCCG 665
QY 661 TTCAATCATGCTGCTGGAATTTGTTTCATCAGATGCTTACCTGCCAGCGTGTCTGTAGA 720
Db 666 TTCAATCATGCTGCTGGAATTTGTTTCATCAGATGCTTACCTGCCAGCGTGTCTGTAGA 725
QY 721 CTCTATTTCTCTGGAGCGTGCCCTCGAGTGCCTGGCATCCCTCTGTCCATGGGCGCCCTG 780
Db 726 CTCTATTTCTCTGGAGCGTGCCCTCGAGTGCCTGGCATCCCTCTGTCCATGGGCGCCCTG 785
QY 781 CACATACCCCTCAGCTTCACCTACATGAACAGATGCTGGATCCCTTGGTGTATTATTTT 840
Db 786 CACATACCCCTCAGCTTCACCTACATGAACAGATGCTGGATCCCTTGGTGTATTATTTT 845
QY 841 TCAAGCCCTCTCTTCCCAAAATTCACAAAGCTCAAAATCTGCAGTCTGAACCCCAAG 900
Db 846 TCAAGCCCTCTCTTCCCAAAATTCACAAAGCTCAAAATCTGCAGTCTGAACCCCAAG 905
QY 901 CAGCCAGGACACTCAAAAACACAAAGCCGGAAGAGATGCCAAATTTGCAACCTCGGTGCG 960
Db 906 CAGCCAGGACACTCAAAAACACAAAGCCGGAAGAGATGCCAAATTTGCAACCTCGGTGCG 965
QY 961 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGCAATGGGATCCC 1020
Db 966 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGCAATGGGATCCC 1025
QY 1021 CACATTCTTGAGTGGCACTGA 1041
Db 1026 CACATTCTTGAGTGGCACTGA 1046
```

```
RESULT 6
AX338373
LOCUS AX338373 1050 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 3 from Patent WO0174904.
ACCESSION AX338373
VERSION AX338373.1 GI:18128870
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
REFERENCE
AUTHORS Majumder,K., Vernet,C.A., Casman,S.J., Wolenc,A.R., Spaderna,S.K.,
Padigaru,M., Mishnu,V.S., Tchernev,V.T., Spytek,K.A., Li,L.,
Baumgartner,J.C. and Gusev,V.Y.
TITLE Novel proteins and nucleic acids encoding same
JOURNAL Patent: WO 0174904-A 3 11-OCT-2001;
FEATURES
Location/Qualifiers
source 1..1050
BASE COUNT 211 a 298 c 279 g 262 t
ORIGIN
Query Match 100.0%; Score 1041; DB 6; Length 1050;
Best Local Similarity 100.0%; Pred. No. 3.5e-249;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTACAACGGGTGCTGCTCCGCATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 60
Db 6 ATGTACAACGGGTGCTGCTCCGCATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 65
QY 61 CTGCTCATTTGGCCCTTTGCTGGGGCACTAGGCAATGGGGTCGCCCTGTGTGTTTC 120
Db 66 CTGCTCATTTGGCCCTTTGCTGGGGCACTAGGCAATGGGGTCGCCCTGTGTGTTTC 125
QY 121 TGCTTCCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTTCAATTTGGCCGTGGCT 180
Db 126 TGCTTCCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTTCAATTTGGCCGTGGCT 185
QY 181 GATTTCCTCTTATGATCTGCCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
Db 186 GATTTCCTCTTATGATCTGCCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 245
QY 241 TGGGCTTTTGGGACATTTCCCTCCGAGTGGGCTCTTACAGTTGGCCATGAACAGGCC 300
Db 246 TGGGCTTTTGGGACATTTCCCTCCGAGTGGGCTCTTACAGTTGGCCATGAACAGGCC 305
QY 301 GGGAGCATCTGTTCCCTTACGGTGGTGGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC 360
Db 306 GGGAGCATCTGTTCCCTTACGGTGGTGGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC 365
QY 361 CACACGCGGTGAACACTATCTCCACCCGGTGGCGCTGGCATGCTCTCCGTCGCAAGAG 420
Db 366 CACACGCGGTGAACACTATCTCCACCCGGTGGCGCTGGCATGCTCTCCGTCGCAAGAG 425
QY 421 GCCTGTCATCTGGGAAAGTGTATCTTTTGTCTGGAGAACCATCTCTCGTGCAGAG 480
Db 426 GCCTGTCATCTGGGAAAGTGTATCTTTTGTCTGGAGAACCATCTCTCGTGCAGAG 485
QY 481 ACGGCCCTCTCTGTGAGAGCTTCATCATGGAGTGGGCCAATGGCTGGCATGACATCATG 540
Db 486 ACGGCCCTCTCTGTGAGAGCTTCATCATGGAGTGGGCCAATGGCTGGCATGACATCATG 545
QY 541 TTCAGCTGAGTCTTTATGCCCTCGGCATCATCTTATTTTGTCTCTTCAAGATTGTT 600
Db 546 TTCAGCTGAGTCTTTATGCCCTCGGCATCATCTTATTTTGTCTCTTCAAGATTGTT 605
QY 601 TGGAGCTGAGGGGAGGAGAGCTGGCCAGACAGGCTGGATGAAGAGGGGACCCCG 660
Db 606 TGGAGCTGAGGGGAGGAGAGCTGGCCAGACAGGCTGGATGAAGAGGGGACCCCG 665
QY 661 TTCAATCATGCTGCTGGAATTTGTTTCATCAGATGCTTACCTGCCAGCGTGTCTGTAGA 720
Db 666 TTCAATCATGCTGCTGGAATTTGTTTCATCAGATGCTTACCTGCCAGCGTGTCTGTAGA 725
QY 721 CTCTATTTCTCTGGAGCGTGCCCTCGAGTGCCTGGCATCCCTCTGTCCATGGGCGCCCTG 780
Db 726 CTCTATTTCTCTGGAGCGTGCCCTCGAGTGCCTGGCATCCCTCTGTCCATGGGCGCCCTG 785
QY 781 CACATACCCCTCAGCTTCACCTACATGAACAGATGCTGGATCCCTTGGTGTATTATTTT 840
Db 786 CACATACCCCTCAGCTTCACCTACATGAACAGATGCTGGATCCCTTGGTGTATTATTTT 845
QY 841 TCAAGCCCTCTCTTCCCAAAATTCACAAAGCTCAAAATCTGCAGTCTGAACCCCAAG 900
Db 846 TCAAGCCCTCTCTTCCCAAAATTCACAAAGCTCAAAATCTGCAGTCTGAACCCCAAG 905
QY 901 CAGCCAGGACACTCAAAAACACAAAGCCGGAAGAGATGCCAAATTTGCAACCTCGGTGCG 960
Db 906 CAGCCAGGACACTCAAAAACACAAAGCCGGAAGAGATGCCAAATTTGCAACCTCGGTGCG 965
QY 961 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGCAATGGGATCCC 1020
Db 966 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGCAATGGGATCCC 1025
QY 1021 CACATTCTTGAGTGGCACTGA 1041
Db 1026 CACATTCTTGAGTGGCACTGA 1046
```

```
Qy 661 TTATCATGTTGGTGGCAATTTGTTTCATCATGCTACCTGCCAGCGTCTCTGCTAGA 720
|||||
Db 666 TTTCATCATGTTGGTGGCAATTTGTTTCATCATGCTACCTGCCAGCGTCTCTGCTAGA 725
|||||
Qy 721 CTCATTTTCCCTCTGGACGGTCCCTCGAGTGCCTGCGATCCCTCTGTCCTATGGGCCCTG 780
|||||
Db 726 CTCATTTTCCCTCTGGACGGTCCCTCGAGTGCCTGCGATCCCTCTGTCCTATGGGCCCTG 785
|||||
Qy 781 CACATAACCCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTCGTGTATTTATTTT 840
|||||
Db 786 CACATAACCCCTCAGCTTCACCTACATGNACAGCATGCTGGATCCCTCGTGTATTTATTTT 845
|||||
Qy 841 TCAAGCCCTCCTTTCCCAAATTTACAAACAGCTCAAAATCTCAGTCTGAAACCCCAAG 900
|||||
Db 846 TCAAGCCCTCCTTTCCCAAATTTACAAACAGCTCAAAATCTCAGTCTGAAACCCCAAG 905
|||||
Qy 901 CAGCAGGACACTCAAAACACAAAGCCGGAAGAGATGCCAATTTGAACTCGGTGCGC 960
|||||
Db 906 CAGCAGGACACTCAAAACACAAAGCCGGAAGAGATGCCAATTTGAACTCGGTGCGC 965
|||||
Qy 961 AGGAGTTGCATCAGTGTGGCAAAATAGTTTCCAAAGCCAGTCTGATGGSCAATGGATCCC 1020
|||||
Db 966 AGGAGTTGCATCAGTGTGGCAAAATAGTTTCCAAAGCCAGTCTGATGGSCAATGGATCCC 1025
|||||
Qy 1021 CACATTGTTGAGTGGCACTGA 1041
|||||
Db 1026 CACATTGTTGAGTGGCACTGA 1046
|||||

RESULT 7
AX305131 1083 bp DNA linear PAT 11-DEC-2001
LOCUS
DEFINITION Sequence 11 from Patent W00187937.
ACCESSION AX305131
VERSION AX305131.1 GI:17644766
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Patterson,C., Lu,D.A., Thornton,M., Lu,Y., Tribouley,C.M.,
Graul,R., Khan,F.A., Gandhi,A.R., Walla,N.K., Nguyen,D.B., Yue,H.,
Hafalia,A., Elliott,V.S., Lal,P., Reddy,R., Kallick,D.A., Tang,T.Y.
and Au-Young,J.
TITLE G-protein coupled receptors
JOURNAL Patent: WO 0187937-A 11 22-NOV-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
1. 1083
/organism="Homo sapiens"
/db_xref="taxon.9606"
/notes="Incyte ID No: 747846Cbl"

BASE COUNT 211 a 314 c 288 g 270 t
ORIGIN

Query Match 100.0%; Score 1041; DB 6; Length 1083;
Best Local Similarity 100.0%; Pred No. 3.5e-249;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTACAAACGGTCTGCTGCCGATCAGAGGGGACACCAATCTCCAGAGTGATGCCGCG 60
|||||
Db 43 ATGTACAAACGGTCTGCTGCCGATCAGAGGGGACACCAATCTCCAGAGTGATGCCGCG 102
|||||
Qy 61 CTGCTCATTTGGCTTTGCTGGGGCAGTACAGGATGGGGTGGCGCTGTGGTTTC 120
|||||
Db 103 CTGCTCATTTGGCTTTGCTGGGGCAGTACAGGATGGGGTGGCGCTGTGGTTTC 162
|||||
Qy 121 TGCTTCCACATGAAGACTGGAAGCCAGCACTGTTTACCTTTTCAATTTGGCCGTGGCT 180
|||||
Db 163 TGCTTCCACATGAAGACTGGAAGCCAGCACTGTTTACCTTTTCAATTTGGCCGTGGCT 222
|||||
Qy 181 GATTTCCTCTTATGATCTGCCTGCCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
|||||
```

```
Db 223 GATTTCCTCTTATGATCTGCCTTTCGGACAGACTATTACCTCAGACGTAGACAC 282
|||||
Qy 241 TGGGCTTTTGGGACATTCCTCGCGAGTGGGCTCTTACGTTGGCCATCAACAGGGCC 300
|||||
Db 283 TGGGCTTTTGGGACATTCCTCGCGAGTGGGCTCTTACGTTGGCCATCAACAGGGCC 342
|||||
Qy 301 GGGAGCATCGTGTTCCTTACGTTGGTGGTGGGAGGTATTTCAAAGTGGTCCACCCC 360
|||||
Db 343 GGGAGCATCGTGTTCCTTACGTTGGTGGTGGGAGGTATTTCAAAGTGGTCCACCCC 402
|||||
Qy 361 CACCACGGGTGAACACTATCTCCACCGGTTGGGCTGGGCTGGCATCGTCTGCACCTGTGG 420
|||||
Db 403 CACCACGGGTGAACACTATCTCCACCGGTTGGGCTGGGCTGGCATCGTCTGCACCTGTGG 462
|||||
Qy 421 GCCCTGTCTATCCTTGGGAACAGTCTATCTTTTGGTGGAGAACCATCTCTGGTGCACAG 480
|||||
Db 463 GCCCTGTCTATCCTTGGGAACAGTCTATCTTTTGGTGGAGAACCATCTCTGGTGCACAG 522
|||||
Qy 481 ACGCCGCTCTCCTGTGAGAGCTTCATCATGAGTGGGCAATGGCTGGCATGACATCATG 540
|||||
Db 523 ACGCCGCTCTCCTGTGAGAGCTTCATCATGAGTGGGCAATGGCTGGCATGACATCATG 582
|||||
Qy 541 TTCAGCTGGAGTTCCTTTATGCCCTCGGCATCATCTTATTTTGTCTCTTCAAGATTGTT 600
|||||
Db 583 TTCAGCTGGAGTTCCTTTATGCCCTCGGCATCATCTTATTTTGTCTCTTCAAGATTGTT 642
|||||
Qy 601 TGGAGCCTGAGCGGAGGACGACGCTGCCAGACAGCTCGGATGMAAGGCCACCCGG 660
|||||
Db 643 TGGAGCCTGAGCGGAGGACGACGCTGCCAGACAGCTCGGATGMAAGGCCACCCGG 702
|||||
Qy 661 TTCATCATGTTGGTGGCAATTTGTTTCATCATGCTACCTGCCAGCGTCTCTGCTAGA 720
|||||
Db 703 TTCATCATGTTGGTGGCAATTTGTTTCATCATGCTACCTGCCAGCGTCTCTGCTAGA 762
|||||
Qy 721 CTCATTTTCCCTCTGGACGGTCCCTCGAGTGCCTGCGATCCCTCTGTCATGGGCCCTG 780
|||||
Db 763 CTCATTTTCCCTCTGGACGGTCCCTCGAGTGCCTGCGATCCCTCTGTCATGGGCCCTG 822
|||||
Qy 781 CACATAACCCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTCGTGTATTTATTTT 840
|||||
Db 823 CACATAACCCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTCGTGTATTTATTTT 882
|||||
Qy 841 TCAAGCCCTCCTTTCCCAAATTTACAAACAGCTCAAAATCTCAGTCTGAAACCCCAAG 900
|||||
Db 883 TCAAGCCCTCCTTTCCCAAATTTACAAACAGCTCAAAATCTCAGTCTGAAACCCCAAG 942
|||||
Qy 901 CAGCCAGGACACTCAAAACACAAAGCCGGAAGAGATGCCAATTTGAACTCGGTGCGC 960
|||||
Db 943 CAGCCAGGACACTCAAAACACAAAGCCGGAAGAGATGCCAATTTGAACTCGGTGCGC 1002
|||||
Qy 961 AGGAGTTGCATCAGTGTGGCAAAATAGTTTCCAAAGCCAGTCTGATGGSCAATGGATCCC 1020
|||||
Db 1003 AGGAGTTGCATCAGTGTGGCAAAATAGTTTCCAAAGCCAGTCTGATGGSCAATGGATCCC 1062
|||||
Qy 1021 CACATTGTTGAGTGGCACTGA 1041
|||||
Db 1063 CACATTGTTGAGTGGCACTGA 1083
|||||

RESULT 8
AX277635 1730 bp DNA linear PAT 01-NOV-2001
LOCUS
DEFINITION Sequence 1 from Patent W00177320.
ACCESSION AX277635
VERSION AX277635.1 GI:16604811
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Xiao,Y.
```


TITLE Regulation of human hm74-like g protein coupled receptor
JOURNAL Patent: WO 0177320-A 1 18-OCT-2001;
Bayer Aktiengesellschaft (DE)
FEATURES
Location/Qualifiers
1..1730
Source /organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 361 a 494 c 444 g 431 t
ORIGIN
Query Match 100.0%; Score 1041; DB 6; Length 1730;
Best Local Similarity 100.0%; Pred. No. 3.6e-249;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTACAACGGGTCGTGCTCCGCGCATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 60
DB 464 ATGTACAACGGGTCGTGCTCCGCGCATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 523
QY 61 CTGCTCATTTGTGGCTTTGTGCTGGGGCACTAGGCAATGGGTGCGCCCTGTGTGGTTTC 120
DB 524 CTGCTCATTTGTGGCTTTGTGCTGGGGCACTAGGCAATGGGTGCGCCCTGTGTGGTTTC 583
QY 121 TGGCTTCACATGAAGACCTCGAAGCCCGACGACTGTTTACCTTTCAATTTGGCCGTGGCT 180
DB 584 TGGCTTCACATGAAGACCTCGAAGCCCGACGACTGTTTACCTTTCAATTTGGCCGTGGCT 643
QY 181 GATTTCCTCTCTATGATCTCGCTCCCTTTTCGGACAGACTATTACCTTCAGACGTAGACAC 240
DB 644 GATTTCCTCTCTATGATCTCGCTCCCTTTTCGGACAGACTATTACCTTCAGACGTAGACAC 703
QY 241 TGGGCTTTTGGGACATTCCTCCGAGTGGGCTCTTTCAGTTGGCCATGAACAGGGCC 300
DB 704 TGGGCTTTTGGGACATTCCTCCGAGTGGGCTCTTTCAGTTGGCCATGAACAGGGCC 763
QY 301 GGAGGATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGATTTTCAAGTGGTCCACCCC 360
DB 764 GGAGGATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGATTTTCAAGTGGTCCACCCC 823
QY 361 CACCACGGGTGAACACTATCTCACCCGGTGGCGCTGGCATCGTCTGCACCCCTGTGG 420
DB 824 CACCACGGGTGAACACTATCTCACCCGGTGGCGCTGGCATCGTCTGCACCCCTGTGG 883
QY 421 GCCTGTCTATCTGGGAACAGTGTATCTTTTGTGTGGAGAACCATCTCTCGTGTGAAGAG 480
DB 884 GCCTGTCTATCTGGGAACAGTGTATCTTTTGTGTGGAGAACCATCTCTCGTGTGAAGAG 943
QY 481 ACGGCGTCTCTGTGAGAGCTTCATCATGAGTGGCCCAATGGCTGGCATGACATCATG 540
DB 944 ACGGCGTCTCTGTGAGAGCTTCATCATGAGTGGCCCAATGGCTGGCATGACATCATG 1003
QY 541 TTCCAGCTGGAGTTCCTTTATGCGCCCTCGGCATCATCTTATTTTGCCTCTTCAAGATTGT 600
DB 1004 TTCCAGCTGGAGTTCCTTTATGCGCCCTCGGCATCATCTTATTTTGCCTCTTCAAGATTGT 1063
QY 601 TGGAGCCTGAGGCGGAGGAGAGCTGGCCAGACAGCTCGGATGAAGAAGGCGACCCCG 660
DB 1064 TGGAGCCTGAGGCGGAGGAGAGCTGGCCAGACAGCTCGGATGAAGAAGGCGACCCCG 1123
QY 661 TTCATCATGGTGGCAATTTGTTTCATCATGCTACCTGCCACCGTGTCTGCTAGA 720
DB 1124 TTCATCATGGTGGCAATTTGTTTCATCATGCTACCTGCCACCGTGTCTGCTAGA 1183
QY 721 CTCTATTTCTCTGGAGGGTGCCTCTGAGTGCCTGGATCCCTCTGTCATGGGCCCTG 780
DB 1184 CTCTATTTCTCTGGAGGGTGCCTCTGAGTGCCTGGATCCCTCTGTCATGGGCCCTG 1243
QY 781 CACATAACCCCTACGCTTCACCTACATGAACAGCATGCTGGATCCCTTGGTGTATTATTT 840
DB 1244 CACATAACCCCTACGCTTCACCTACATGAACAGCATGCTGGATCCCTTGGTGTATTATTT 1303
QY 841 TCAAGCCCTCTCTTCCCAAAATTTCTACACAGCTCAAAATCTGCAGTCTGAACCCAG 900
DB 1304 TCAAGCCCTCTCTTCCCAAAATTTCTACACAGCTCAAAATCTGCAGTCTGAACCCAG 1363

QY 901 CAGCAGGACACTCAAAACACAAAGCCGGAAGAGATGCAATTTTCGAACCTCGGTGCG 960
DB 1364 CAGCAGGACACTCAAAACACAAAGCCGGAAGAGATGCAATTTTCGAACCTCGGTGCG 1423
QY 961 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGCAATGGGATCCC 1020
DB 1424 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGCAATGGGATCCC 1483
QY 1021 CACATTGTTGAGTGGCACTGA 1041
DB 1484 CACATTGTTGAGTGGCACTGA 1504
RESULT 9
AX299707
LOCUS AX299707 2331 bp DNA linear PAT 26-NOV-2001
DEFINITION Sequence 3 from Patent WO0173029.
ACCESSION AX299707
VERSION AX299707.1 GI:17129251
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Ye,J.C., Cravchik,A.C., di Francesco,V.C. and Beasley,E.M.
TITLE Isolated human g-protein coupled receptors, nucleic acid molecules
encoding human gpcr proteins, and uses thereof
JOURNAL Patent: WO 0173029-A 3 04-OCT-2001;
PE Corporation (NY) (US)
FEATURES
Location/Qualifiers
Source 1..2331
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 497 a 620 c 592 g 622 t
ORIGIN
Query Match 100.0%; Score 1041; DB 6; Length 2331;
Best Local Similarity 100.0%; Pred. No. 3.7e-249;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTACAACGGGTCGTGCTCCGCGCATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 60
DB 311 ATGTACAACGGGTCGTGCTCCGCGCATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 370
QY 61 CTGCTCATTTGTGGCTTTGTGCTGGGGCACTAGGCAATGGGTGCGCCCTGTGTGGTTTC 120
DB 371 CTGCTCATTTGTGGCTTTGTGCTGGGGCACTAGGCAATGGGTGCGCCCTGTGTGGTTTC 430
QY 121 TGGCTTCACATGAAGACCTGGAAGCCCGACGACTGTTTACCTTTTCAATTTGGCGGTGGCT 180
DB 431 TGGCTTCACATGAAGACCTGGAAGCCCGACGACTGTTTACCTTTTCAATTTGGCGGTGGCT 490
QY 181 GATTTCCTCTCTATGATCTCGCTCCCTTTTCGGACAGACTATTACCTTCAGACGTAGACAC 240
DB 491 GATTTCCTCTCTATGATCTCGCTCCCTTTTCGGACAGACTATTACCTTCAGACGTAGACAC 550
QY 241 TGGGCTTTTGGGACATTCCTCCGAGTGGGCTCTTTCAGTTGGCCATGAACAGGGCC 300
DB 551 TGGGCTTTTGGGACATTCCTCCGAGTGGGCTCTTTCAGTTGGCCATGAACAGGGCC 610
QY 301 GGAGGATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGATTTTCAAGTGGTCCACCCC 360
DB 611 GGAGGATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGATTTTCAAGTGGTCCACCCC 670
QY 361 CACCACGGGTGAACACTATCTCCACCCGGTGGCGCTGGCATCGTCTGCACCCCTGTGG 420
DB 671 CACCACGGGTGAACACTATCTCCACCCGGTGGCGCTGGCATCGTCTGCACCCCTGTGG 730
QY 421 GCCTGTCTATCTGGGAACAGTGTATCTTTTGTGTGGAGAACCATCTCTGCGTGTGAAGAG 480
DB 731 GCCTGTCTATCTGGGAACAGTGTATCTTTTGTGTGGAGAACCATCTCTGCGTGTGAAGAG 790

```

Qy 481 ACGCCGCTCTCTGTGAGAGCTTCATCATGAGTCGCGCAATGGCTGGCATGACATCATG 540
Db 791 ACGCCGCTCTCTGTGAGAGCTTCATCATGAGTCGCGCAATGGCTGGCATGACATCATG 850
Qy 541 TTCAGCTGGAGTCTTTATGCCCCCTCGGCATCATCTTATTTTGGCTCTCTCAAGATTGTT 600
Db 851 TTCAGCTGGAGTCTTTATGCCCCCTCGGCATCATCTTATTTTGGCTCTCTCAAGATTGTT 910
Qy 601 TGGAGCTGAGCGGAGCGACAGCTGCGCCAGACAGCTCGGATGAAGAAGCGCCACCGG 660
Db 911 TGGAGCTGAGCGGAGCGACAGCTGCGCCAGACAGCTCGGATGAAGAAGCGCCACCGG 970
Qy 661 TTCATCATGTGTGGCAATTGTTCATCATCATGCTACCTGCCAGCGTGTCTGCTGAGA 720
Db 971 TTCATCATGTGTGGCAATTGTTCATCATCATGCTACCTGCCAGCGTGTCTGCTGAGA 1030
Qy 721 CTCATTTCTCTGGAGGTCCTCGAGTCGCTGCGATCCCTCTGTCATGGGGCCCTG 780
Db 1031 CTCATTTCTCTGGAGGTCCTCGAGTCGCTGCGATCCCTCTGTCATGGGGCCCTG 1090
Qy 781 CACATAACCTTCAGCTACCTACATGAACAGCATGCTGGATCCCTGGTGTATTTT 840
Db 1091 CACATAACCTTCAGCTACCTACATGAACAGCATGCTGGATCCCTGGTGTATTTT 1150
Qy 841 TCAAGCCCTCTCTTCCCAATTTACAACAAGCTCAAAATCTCAGTCTGAAACCCCAAG 900
Db 1151 TCAAGCCCTCTCTTCCCAATTTACAACAAGCTCAAAATCTCAGTCTGAAACCCCAAG 1210
Qy 901 CAGCAGGACACTCAAAACACAAAGCGGAGAGATGCCAATTCGAACCTCGGTGCG 960
Db 1211 CAGCAGGACACTCAAAACACAAAGCGGAGAGATGCCAATTCGAACCTCGGTGCG 1270
Qy 961 AGGAGTTCATCAGTGGCAATAGTTTCCAAAGCAGTCTGATGGCAATGGATGCC 1020
Db 1271 AGGAGTTCATCAGTGGCAATAGTTTCCAAAGCAGTCTGATGGCAATGGATGCC 1330
Qy 1021 CACATTTGTGAGTGGCACTGA 1041
Db 1331 CACATTTGTGAGTGGCACTGA 1351

RESULT 10
AC026331
LOCUS Homo sapiens chromosome 12 clone RP11-507N20, WORKING DRAFT
DEFINITION
SEQUENCE, 10 unordered pieces.
ACCESSION AC026331
VERSION AC026331.18 GI:14547388
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 179172)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.D., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,J., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,X., Johnson,R., Jollivet,S.,

```

```

Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtege,O., Lieu,C., Liu,J., Liu,W.,
Loulsegh,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokwkw,S.,
Oruh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,
Ruiz,S., Saverly,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
Sisson,I., Sodergren,E., Sonaika,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 179172)
Worley,K.C.
Direct Submission
Submitted (22-MAR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jun 25, 2001 this sequence version replaced gi:13877178.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: HAQ1
Center clone name: RP11-507N20
----- Summary Statistics
Sequencing vector: Plasmid: M77789
Sequencing vector: M13: L08821
Chemistry: Dye-primer Bodipy: 31% of reads
Chemistry: Dye-terminator Big Dye: 69% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 175811 bases at least Q40
Consensus quality: 178545 bases at least Q30
Consensus quality: 179678 bases at least Q20
Estimated insert size: 176741; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-gel estimation
Quality coverage: 9.4x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 78354: contig of 78354 bp in length
* 78355 78454: gap of unknown length
* 78455 104278: contig of 25824 bp in length
* 104279 104378: gap of unknown length
* 104379 125803: contig of 21425 bp in length
* 125804 125903: gap of unknown length
* 125904 145254: contig of 19351 bp in length
* 145255 145354: gap of unknown length
* 145355 153279: contig of 7925 bp in length
* 153280 153379: gap of unknown length
* 153380 160977: contig of 7598 bp in length
* 160978 161077: gap of unknown length

```

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT


```
|||||
Db 16591 CACATTGTTGAGTGGCACTGA 16611
|||||
RESULT 12
AX338374
LOCUS
DEFINITION
ACCESSION AX338374
VERSION AX338374.1 GI:18128871
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (sites)
JOURNAL Majumder,K., Vernet,C.A., Casman,S.J., Wolenc,A.R., Spaderna,S.K.,
Novel proteins and nucleic acids encoding same Padigaru,M., Mishnu,V.S., Tchernev,V.T., Spytek,K.A., Li,L.,
Patent: WO 0174904-A 4 11-OCT-2001; Baumgartner,J.C. and Gusev,V.Y.
FEATURES
Location/Qualifiers
SOURCE i. .1104
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 216 a 318 c 295 g 275 t
ORIGIN
Query Match 99.7%; Score 1037.8; DB 6; Length 1104;
Best Local Similarity 99.8%; Pred. No. 2.2e-248;
Matches 1039; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGTACAAGGGTCGTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATGCCGCG 60
Db 60 ATGTACAAGGGTCGTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATGCCGCG 119
QY 61 CTGCTCATTTGTGGCCCTTTGTGCTGGCGCACTAGGCAATGGGTCGCGCTGTGTGGTTTC 120
Db 120 CTGCTCATTTGTGGCCCTTTGTGCTGGCGCACTAGGCAATGGGTCGCGCTGTGTGGTTTC 179
QY 121 TGGTTCACATGAAGACCTGGAGCCGACGACTGTTTACCTTTTCAATTTGGCCGTGGCT 180
Db 180 TGGTTCACATGAAGACCTGGAGCCGACGACTGTTTACCTTTTCAATTTGGCCGTGGCT 239
QY 181 GATTTCTCTTATGATCTGCCTGCTTTTCGACAGACTATTACCTCAGACGTAGACAC 240
Db 240 GATTTCTCTTATGATCTGCCTGCTTTTCGACAGACTATTACCTCAGACGTAGACAC 299
QY 241 TGGGCTTTTGGGACATTCCTCCGAGTGGGCTCTTACGTTGGCCATGAACAGGCGC 300
Db 300 TGGGCTTTTGGGACATTCCTCCGAGTGGGCTCTTACGTTGGCCATGAACAGGCGC 359
QY 301 GGGAGCATGTTTCCTTACGGTGGTGGCTGCGGACAGTATTTCAAGTGTGTCACCCC 360
Db 360 GGGAGCATGTTTCCTTACGGTGGTGGCTGCGGACAGTATTTCAAGTGTGTCACCCC 419
QY 361 CACCAGCGGTGAACACTATCTCCACCGCGGTGGCGCTGGCATCGTCTGCACCCCTGTG 420
Db 420 CACCAGCGGTGAACACTATCTCCACCGCGGTGGCGCTGGCATCGTCTGCACCCCTGTG 479
QY 421 GCCTGTGTCATCTCGGGAACAGTGTATCTTTTCTGGAGAACCATCTCTGCGTGAAGAG 480
Db 480 GCCTGTGTCATCTCGGGAACAGTGTATCTTTTCTGGAGAACCATCTCTGCGTGAAGAG 539
QY 481 ACGGCGCTCTCTGTGAGAGCTTCATCATGGAGTCGGCCAAATGGCTGGCATGACATG 540
Db 540 ACGGCGCTCTCTGTGAGAGCTTCATCATGGAGTCGGCCAAATGGCTGGCATGACATG 599
QY 541 TTCCAGCTGGAGTCTTTATGCCCCCTCGGCATCATCTTATTTTGTCTTCAAGATTGTT 600
Db 600 TTCCAGCTGGAGTCTTTATGCCCCCTCGGCATCATCTTATTTTGTCTTCAAGATTGTT 659

|||||
Db 601 TGGAGCCTGAGCGGAGGAGCAGAGCTGGCAGACAGAGCTCGGATGAAGAGGCGACCCGG 660
Db 660 TGGAGCCTGAGCGGAGGAGCAGAGCTGGCAGACAGAGCTCGGATGAAGAGGCGACCCGG 719
QY 661 TTCATCATGTTGTTGGCAATTTGTTTCATCATCATGCTACTGCCACAGCGTGTCTGTCTAGA 720
Db 720 TTCATCATGTTGTTGGCAATTTGTTTCATCATCATGCTACTGCCACAGCGTGTCTGTCTAGA 779
QY 721 CTCATTTCCTCTGAGCGGTGCCCTCGAGTGCCTCGATGCCATCCCTCTGTCCATGSGGCCCTG 780
Db 780 CTCATTTCCTCTGAGCGGTGCCCTCGAGTGCCTCGATGCCATCCCTCTGTCCATGSGGCCCTG 839
QY 781 CACATAACCTCAGCTTACCTACATGAACAGCATGCTGATCCCTCGCTGTGTATTTT 840
Db 840 CACATAACCTCAGCTTACCTACATGAACAGCATGCTGATCCCTCGCTGTGTATTTT 899
QY 841 TCAAGCCCTCTCTTCCCAAAATTTCTACACAGCTCAAAATCTGCAAGTGTGAACCCCAAG 900
Db 900 TCAAGCCCTCTCTTCCCAAAATTTCTACACAGCTCAAAATCTGCAAGTGTGAACCCCAAG 959
QY 901 CAGCCAGGACACTCAAAACACAAAGCCGGAAGAGATGCCAATTTTCGAACCTCGGTGCG 960
Db 960 CAGCCAGGACACTCAAAACACAAAGCCGGAAGAGATGCCAATTTTCGAACCTCGGTGCG 1019
QY 961 AGGAGTTGCATCAGTGTGCAAAATAGTTTCCAAAGCCAGTCTGATGGCAATGGGATCCC 1020
Db 1020 AGGAGTTGCATCAGTGTGCAAAATAGTTTCCAAAGCCAGTCTGATGGCAATGGGATCCC 1079
QY 1021 CACATTGTTGAGTGGCACTGA 1041
Db 1080 CACATTGTTGAGTGGCACTGA 1100

RESULT 13
AX147766
LOCUS
DEFINITION
ACCESSION AX147766
VERSION AX147766.1 GI:14346804
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 888)
JOURNAL Vogeli,G., Wood,L.S., Parodi,L.A., Hiebsch,R.R., Lind,P.,
Novel g protein-coupled receptors Slightom,J., Schellin,K.A., Kaytes,P.S., Bannigan,C.M., Ruff,V.,
Patent: WO 0136473-A 11 25-MAY-2001; Sejlitz,T. and Huff,R.M.
FEATURES
Location/Qualifiers
SOURCE i. .888
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 174 a 254 c 231 g 229 t
ORIGIN
Query Match 85.3%; Score 888; DB 6; Length 888;
Best Local Similarity 100.0%; Pred. No. 5.3e-211;
Matches 888; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 CTGCTCATTTGTGCCCTTTGTGCTGGCGCACTAGGCAATGGGTCGCGCTGTGTGGTTTC 120
Db 1 CTGCTCATTTGTGCCCTTTGTGCTGGCGCACTAGGCAATGGGTCGCGCTGTGTGGTTTC 60
QY 121 TGTCTCCACATGAAGACCTGGAAAGCCGACGACTGTTTACCTTTTCAATTTGGCCGTGGCT 180
Db 61 TGTCTCCACATGAAGACCTGGAAAGCCGACGACTGTTTACCTTTTCAATTTGGCCGTGGCT 120
QY 181 GATTTCTCTTATGATCTGCCTGCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
Db 121 GATTTCTCTTATGATCTGCCTGCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 180
```


DEFINITION Sequence 6104 from Patent WO0194629.

ACCESSION AX335595

VERSION AX335595.1 GI:18126314

KEYWORDS human.

SOURCE

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (sites)

AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,

Horrigan, S., Soppet, D.R. and Weaver, Z.

TITLE Cancer gene determination and therapeutic screening using signature

gene sets

JOURNAL Patent: WO 0194629-A 6104 13-DEC-2001;

Avalon Pharmaceuticals (US)

FEATURES Location/Qualifiers

source 1..2051

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 470 a 530 c 525 g 526 t

ORIGIN

Query Match 35.3%; Score 367.6; DB 6; Length 2051;

Best Local Similarity 63.9%; Pred. No. 4.5e-81;

Matches 597; Conservative 0; Mismatches 319; Indels 18; Gaps 2;

QY 1 ATGTACAACGGGTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 60

DB 97 ATAGACAAGAAGACTGCTGTGTCGAGATGACTTATGCCAAGGTGTGCCGCCG 156

QY 61 CTGCTCATTGTGCCCTTTGTCTGGGGCAGTGGGCAATGGGTCGCCCTGTGTGTTT 120

DB 157 GTGTGGGGCTGGAGTTTATCTTTGGGCTTCTGGCAATGGCCTTGCCCTGTGATTTT 216

QY 121 TGCCTCCACATGAAGACCTGGAGCCAGCAGCTGTTTACCTTTTCAATTTGGCCGTGGCT 180

DB 217 TGTTCACCTCAAGCTCTGGAATCCAGCCGATTTTCTGTGTCAACCTGGCAGTAGCT 276

QY 181 GATTTCTCTCTATGATCTGCCCTTCTGGACAGACTATTACCTCAGACGTAGACAC 240

DB 277 GACTTCTACTGATCATCTGCCCTGCGGTGCTGATGGACTACTATGTGCGCGGTTCAGAC 336

QY 241 TGGGCTTTTGGGACATTCCTCCGAGTGGGCTCTTACATGTGGCCATGAACAGGCC 300

DB 337 TGGAACTTTGGGACATCCCTTGGCGCTGGTCTCTTCATGTTTGCCATGAACGCCAG 396

QY 301 GGGAGCATCTGTTCTTACGGTGGTGGCTGGCGACAGGTATTTCAAAGTGGTCCACCC 360

DB 397 GGCAGCATCATCTTCTCACGGTGGTGGCGGTAGACAGGTATTTCCGGGTGGTCCATCC 456

QY 361 CACACGCGTGAACACTATCTCCACCGGGTGGCGGCTGGCATCGTCTGCCCTGTGG 420

DB 457 CACACGCGCTGAACAGATCTCCAATTGGACAGAGCCATCATCTCTGCCCTCTGTGG 516

QY 421 GCCTGGTCACTCTGGGAACAGTGTATCTTTTGTGGGAACCATCTCTCGGTGCAAGAG 480

DB 517 GGCATCACTCTTGCCCTTAACAGTCCACCTCTTGAAGAAGATTGCTGTATCCAGATGGC 576

QY 481 ACGGCCGTCTCTGTGAGACTTCATCATGGAGTGGCCAAATGGCTGGCATGACATCATG 540

DB 577 CCTGCAAAATGTGTCATCAGCTCTCAGCATCTGCCATCTGCCCTGGCGACGAAGTATG 636

QY 541 TTCAGCTGGAGTTCTTTATGCCCTGGCATCATCTTATTTTGTCTCTTCAAGATTGTT 600

DB 637 TTCTCTGGAGTTCTCTTGGCCCTGGGCATCATCTGTTCTGCTCAGCCAGAAATTATC 696

QY 601 TGGAGCTGAGGGGAGGAGAGCTGGCCAGACAGGCTCGGATGAAGAAGGCGACCCGG 660

DB 697 TGGAGCTGGCGCAGAG- --ACAAATGGACCGCATGTCCAAGATCAAGAGGCCATCAC 753

QY 661 TTCAATCATGTGTGGCAATTGTGTCATCACAATGCTACCTGCCCGACGCTGTCTGTCTAGA 720

DB 754 TTCAATCATGTGTGGCCATCGTCTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 813

QY 721 CTCTATTTCCTCTGGACGGTGCCTC-----GAGTGCCTCGCATCCCTCT 765

DB 814 ATCCGCATCTTCTGGCTCTCCGCACACTTCGGGCACGAGAAATGTGAAGTGTACCGCTCG 873

QY 766 GTCCATGGGGCCCTGCACATAACCCCTCAGCTTACCTACATGAACAGCATGCTGGATCCC 825

DB 874 GTGGACCTGGCGTTCTTTATCACTCTCAGCTTCACTACATGAACAGCATGCTGGACCCC 933

QY 826 CTGGTGTATATTTTCAAGCCCTCTTCCCAAAATTTCTACAACAGCTCAAAATCTGC 885

DB 934 GTGGTGTACTACTTCTCCAGCCCATCTTCCCAACTTCTTCTCCACTTTTGATCAACCCG 993

QY 886 AGTCTGAAACCCCAAGCAGCAGGACACTCAAAAA 919

DB 994 TGCTCCAGAGGAGATGACAGGTGAGCCAGATA 1027

Search completed: October 29, 2002, 03:17:28

Job time : 2068 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2002, 02:41:47 ; Search time 208 Seconds
(without alignments)
8592.827 Million cell updates/sec

Title: US-09-886-041-1
Perfect score: 1041
Sequence: 1 atgtacaacgggtgctgctg.....acattgtgagtggaactga 1041

Scoring table: OLIGO.NUC
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_032802.*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1041	100.0	1041	22 AAS12581	CDNA encoding nove
2	1041	100.0	1041	22 AAS07946	Human CDNA encodin
3	1041	100.0	1041	22 AAS1008	Human gPCR11 codi
4	1041	100.0	1050	22 ABA81529	Human GPCR1a polyn
5	1041	100.0	1050	22 ABA81530	Human GPCR1b polyn
6	1041	100.0	1083	24 AAD26371	Human G-protein co
7	1041	100.0	1730	24 AAS18501	CDNA encoding HM74
8	1041	100.0	2331	22 AAS12582	Gene encoding nove
9	939	90.2	1104	22 ABA81531	Human GPCR1c polyn

10	888	85.3	888	22	AAH50974	Human NGPCRL1 codi
11	32	3.1	1092	22	AAS07952	Human cDNA encodin
12	32	3.1	1164	21	AAA30658	Human G protein-co
13	32	3.1	1164	21	AAA30738	DNA encoding human
14	32	3.1	1361	20	AAAX1671	G-protein coupled
15	24	2.3	24	22	AAH51053	Oligonucleotide SE
16	24	2.3	24	22	AAH51082	Human NGPCRL1 PCR
17	24	2.3	52	22	AAH51050	Oligonucleotide SE
18	24	2.3	52	22	AAH51079	Human NGPCRL1 PCR
19	23	2.2	33	22	AAH51029	Human NGPCRL1 PCR
20	21	2.0	33	22	AAH51028	Human NGPCRL1 PCR
21	20	1.9	20	22	AAH51030	Human NGPCRL1 PCR
22	20	1.9	20	22	AAH51031	Human NGPCRL1 PCR
23	20	1.9	20	22	AAH51051	Oligonucleotide SE
24	20	1.9	20	22	AAH51080	Human NGPCRL1 PCR
25	20	1.9	48	22	AAH51052	Oligonucleotide SE
26	20	1.9	48	22	AAH51081	Human NGPCRL1 PCR
27	20	1.9	4715	22	AAAL37148	Human musculoskele
28	19	1.8	912	23	ABL17061	Drosophila melanog
29	19	1.8	2947	23	ABL02109	Drosophila melanog
30	19	1.8	2977	23	ABL17060	Drosophila melanog
31	19	1.8	3482	23	ABL19418	Drosophila melanog
32	19	1.8	10098	22	AAS27822	DNA encoding novel
33	19	1.8	12619	22	AAH26493	Human low density
34	19	1.8	19175	23	ABL02108	Drosophila melanog
35	18	1.7	433	19	AAV09268	Nucleotide sequenc
36	18	1.7	433	19	AAV04273	Secreted protein C
37	18	1.7	444	22	AAI92474	Human polynucleoti
38	18	1.7	451	20	AAV86048	EST clone C195. H
39	18	1.7	464	22	AAI82739	Human polynucleoti
40	18	1.7	477	22	AAS36762	Human cardiovascular
41	18	1.7	477	22	AAS36763	Human cardiovascular
42	18	1.7	489	22	AAS35641	Human cardiovascular
43	18	1.7	733	19	AAV04274	Secreted protein C
44	18	1.7	757	22	AAI96369	Human neuroblastom
45	18	1.7	767	21	AAA81592	N. meningitidis pa

ALIGNMENTS

RESULT 1

AAS12581

ID AAS12581 standard; cDNA; 1041 BP.

XX AAS12581;

DT 19-DEC-2001 (first entry)

XX cDNA encoding novel human G protein-coupled receptor (GPCR).

DE Human; G-protein coupled receptor; GPCR; chemokine receptor; protease;
KW hyperproliferative disorder; neurological disorder; psychiatric disease;
KW inflammatory disorder; respiratory disorder; gene therapy; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..1041

FT /tag= a

FT /product= "GPCR protein"

XX WO200173029-A2.

XX 04-OCT-2001.

XX 27-MAR-2001; 2001WO-US09522.

XX 27-MAR-2000; 2000US-192419P.

XX 06-SEP-2000; 2000US-230459P.

XX 20-SEP-2000; 2000US-0666535.

XX (PEKE) PE CORP NY.

XX Ye J, Cravchik A, Di Francesco V, Beasley EM;
XX WPI: 2001-616503/71.
XX P-PSDB: RAU06197.
XX Novel human G-protein coupled receptor proteins and nucleic acid
XX molecules encoding the protein for use in developing human therapeutics
XX and diagnostic compositions and for identifying modulators of the
XX protein -
XX
XX Claim 23; Fig 1; 66pp; English.
XX The present invention relates to the isolation of a novel human G-protein
XX coupled receptor (GPCR) which is related to the chemokine receptor
XX subfamily. The cDNA and gene sequences encoding for GPCR are also
XX given in the invention. The sequences of the invention are useful
XX for diagnosing and treating diseases or conditions mediated by human
XX proteases. Such diseases include hyperproliferative disorders
XX (e.g. hyperplasia), neurological disorders (e.g. Parkinson's disease),
XX psychiatric diseases (e.g. schizophrenia), inflammatory disorders
XX (e.g. diabetes) and respiratory disorders (e.g. adult respiratory
XX distress syndrome, ARDS). The GPCR protein is also useful for identifying
XX a modulator of the expression of the protein. It also serves as a target
XX for identifying agents for use in mammalian therapeutic applications,
XX e.g. a human drug, particularly modulating a biological or pathological
XX response in a cell or tissue that expresses the protein, in biological
XX assays related to GPCRs that are related to members of the chemokine
XX receptor subfamily, in drug screening assays and in competition binding
XX assays. GPCR is also useful in diagnosing a disease or predisposition to
XX a disease mediated by the peptide, in pharmacogenomic analysis. The
XX polynucleotide sequences can also be used in gene therapy. The present
XX sequence encodes for the novel human GPCR of the invention.
XX
XX Sequence 1041 BP; 208 A; 294 C; 278 G; 261 T; 0 other;

Query Match 100.0%; Score 1041; DB 22; Length 1041;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTACACGGTCTGCTGCGGATCGAGGGGACACCATCTCCAGGTGATGCCGCG 60
Db 1 ATGTACACGGTCTGCTGCGGATCGAGGGGACACCATCTCCAGGTGATGCCGCG 60
Qy 61 CTGCTCATTTGGGCTTTGTGCTGGGCGACTAGGCAATGGGGTGGCCCTGTGGTTTC 120
Db 61 CTGCTCATTTGGGCTTTGTGCTGGGCGACTAGGCAATGGGGTGGCCCTGTGGTTTC 120
Qy 121 TGCTTCCACATGAAGACCTGGAAGCCAGCAGTGTTCACCTTTTCAATTTGGCGTGGCT 180
Db 121 TGCTTCCACATGAAGACCTGGAAGCCAGCAGTGTTCACCTTTTCAATTTGGCGTGGCT 180
Qy 181 GATTTCCTCCTTATGATCTGCTGCTTTTCGGACAGACTATTACCTTCAGACGTAGACAC 240
Db 181 GATTTCCTCCTTATGATCTGCTGCTTTTCGGACAGACTATTACCTTCAGACGTAGACAC 240
Qy 241 TGGGCTTTTGGGACATTCCTTCCGAGTGGGGCTCTTACGTTGGCCATGAACAGGGCC 300
Db 241 TGGGCTTTTGGGACATTCCTTCCGAGTGGGGCTCTTACGTTGGCCATGAACAGGGCC 300
Qy 301 GGGAGACATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTTCAAGTGGTCCACCC 360
Db 301 GGGAGACATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTTCAAGTGGTCCACCC 360
Qy 361 CACCACGGGTGAACACTATCTCACCAGGTGGCGGTGGCATCTGCGACCTGTGG 420
Db 361 CACCACGGGTGAACACTATCTCACCAGGTGGCGGTGGCATCTGCGACCTGTGG 420
Qy 421 GCCCTGGTTCATCCTGGGAACAGTGTATCTTTTGTGGAGAACCACTCTCTGGTGAAGAG 480
Db 421 GCCCTGGTTCATCCTGGGAACAGTGTATCTTTTGTGGAGAACCACTCTCTGGTGAAGAG 480
Qy 481 ACGGCGCTCTCTGTGAGAGGCTTTCATCATGAGAGTCGGCCAAATGGCTGGCATCATG 540

Db 481 ACGGCGCTCTCTGTGAGAGGCTTTCATCATGAGAGTCGGCCAAATGGCTGGCATCATG 540
Qy 541 TTCAGCTGGAGTTCCTTATGCGGCTGGCATCATCTTATTTTCTCTTCAAGATGTT 600
Db 541 TTCAGCTGGAGTTCCTTATGCGGCTGGCATCATCTTATTTTCTCTTCAAGATGTT 600
Qy 601 TGGAGCCTTGGGCGAGGACAGCTGGCCAGACAGGCTCGGATGAAGAGGCGACCCGG 660
Db 601 TGGAGCCTTGGGCGAGGACAGCTGGCCAGACAGGCTCGGATGAAGAGGCGACCCGG 660
Qy 661 TTCATCATGTTGGGCAATTTGTTCATCATGCTTACCTGCCCCAGGCTCTGCTAGA 720
Db 661 TTCATCATGTTGGGCAATTTGTTCATCATGCTTACCTGCCCCAGGCTCTGCTAGA 720
Qy 721 CTCATTTTCTCTCGAGGCTGCTCGAGTCCCTCGGATCCCTCTGTCCATGGGCGCTG 780
Db 721 CTCATTTTCTCTCGAGGCTGCTCGAGTCCCTCGGATCCCTCTGTCCATGGGCGCTG 780
Qy 781 CACATAACCCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTGGTGTATTTT 840
Db 781 CACATAACCCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTGGTGTATTTT 840
Qy 841 TCAGGCGCTCTCTTCCCAATTTCTACAAAGCTCAAAATCTGAGTCTGAAACCCAG 900
Db 841 TCAGGCGCTCTCTTCCCAATTTCTACAAAGCTCAAAATCTGAGTCTGAAACCCAG 900
Qy 901 CAGCCAGGACACTCAAAACACAAAGGCGGAGAGATGCCAATTTGGAACCTCGGTGC 960
Db 901 CAGCCAGGACACTCAAAACACAAAGGCGGAGAGATGCCAATTTGGAACCTCGGTGC 960
Qy 961 AGGAGTTGTCATGAGTGGCAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGATCCC 1020
Db 961 AGGAGTTGTCATGAGTGGCAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGATCCC 1020
Qy 1021 CACATTGTTGAGTGGCACTGA 1041
Db 1021 CACATTGTTGAGTGGCACTGA 1041
RESULT 2
AAS07946
ID AAS07946 standard; cDNA; 1041 BP.
XX AAS07946;
AC AC
XX AC
DT 23-OCT-2001 (first entry)
XX
DE Human cDNA encoding G-protein coupled receptor, hRUP19.
XX
KW Human; G-protein coupled receptor; GPCR; hRUP19; agonist;
KW Inverse agonist; lung cancer; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1041
FT /tag= a
FT /product= "hRUP19"
XX
PN WO200136471-A2.
XX
XX 25-MAY-2001.
XX
XX 16-NOV-2000; 2000WO-US31509.
XX
XX 17-NOV-1999; 99US-0166088.
PR 17-NOV-1999; 99US-0166099.
PR 17-NOV-1999; 99US-0166369.
PR 23-DEC-1999; 99US-0171900.
PR 23-DEC-1999; 99US-0171901.
PR 23-DEC-1999; 99US-0171902.
PR 11-FEB-2000; 2000US-0181749.

PR 14-MAR-2000; 2000US-0189258.
PR 14-MAR-2000; 2000US-0189259.
PR 10-APR-2000; 2000US-0195898.
PR 10-APR-2000; 2000US-0195899.
PR 10-APR-2000; 2000US-0196078.
PR 28-APR-2000; 2000US-0200419.
PR 12-MAY-2000; 2000US-0203630.
PR 12-JUN-2000; 2000US-0210741.
PR 21-AUG-2000; 2000US-0210982.
PR 21-AUG-2000; 2000US-0226760.
PR 26-SEP-2000; 2000US-0235418.
PR 26-SEP-2000; 2000US-0235779.
PR 20-OCT-2000; 2000US-0242332.
PR 20-OCT-2000; 2000US-0242343.
XX (AREN-) ARENA PHARM INC.
XX
XX Chen R, Dang HT, Lowitz KP;
PI
XX WPI; 2001-355616/37.
DR P-PSDB; AAU04373.
DR
XX Endogenous and non-endogenous versions of human G-protein coupled
PT receptors for direct identification of candidate compounds as agonists,
PT inverse agonists or partial agonists for use as therapeutic agents -
XX
XX Claim 47; Page 110; 159pp; English.
XX
XX The sequence encodes a human G-protein coupled receptor (GPCR),
CC hRUP19 The endogenous and non-endogenous, constitutively activated
CC versions of human G-protein coupled receptors (GPCR), are useful for
CC direct identification of candidate compounds as receptor agonists,
CC inverse agonists or partial agonists having applicability as therapeutic
CC agents for treating diseases related to GPCR, e.g. lung cancer.
CC Non-endogenous version of human GPCRs are also utilized in research
CC settings and in vitro and in vivo system, incorporating GPCRs can be
CC utilised to elucidate and understand the roles these receptors
CC play in the human condition, both normal and diseased.
XX
XX Sequence 1041 BP; 208 A; 294 C; 278 G; 261 T; 0 other;
SQ

Query Match 100.0%; Score 1041; DB 22; Length 1041;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTACAACGGGTGCTGTCGCGCATCGAGGGGACACCATCTCCAGGTGATCGCGCG 60
DB 1 ATGTACAACGGGTGCTGTCGCGCATCGAGGGGACACCATCTCCAGGTGATCGCGCG 60
QY 61 CTGCTCATTTGGCCCTTTGCTGGCGCAGCTAGGCAATGGGCTGCTGTTTC 120
DB 61 CTGCTCATTTGGCCCTTTGCTGGCGCAGCTAGGCAATGGGCTGCTGTTTC 120
QY 121 TCGTTCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTTCAATTTGGCGTGGCT 180
DB 121 TCGTTCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTTCAATTTGGCGTGGCT 180
QY 181 GATTTCCTCTTATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 181 GATTTCCTCTTATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 241 TGGGCTTTTGGGACATTTCCCTGCGGAGTGGGCTCTTCCAGCTTGGCCATGAACAGGGCC 300
DB 241 TGGGCTTTTGGGACATTTCCCTGCGGAGTGGGCTCTTCCAGCTTGGCCATGAACAGGGCC 300
QY 301 GGGAGCATCGTGTTCCTTACGTTGGTGGCTGCGGACAGTATTTCAAAGTGTTCACCCC 360
DB 301 GGGAGCATCGTGTTCCTTACGTTGGTGGCTGCGGACAGTATTTCAAAGTGTTCACCCC 360
QY 361 CACCAGCGGTGAACACTATATCCACCCGGGTGGCGGCTGGCATCGTCTGCACCCCTGTGG 420
DB 361 CACCAGCGGTGAACACTATATCCACCCGGGTGGCGGCTGGCATCGTCTGCACCCCTGTGG 420

QY 421 GCCCTGGTCATCTCTGGGAACAGTGTATCTTTTCTGGAGAACCATCTCTCTGCGTGAAGAG 480
DB 421 GCCCTGGTCATCTCTGGGAACAGTGTATCTTTTCTGGAGAACCATCTCTCTGCGTGAAGAG 480
QY 481 ACGGCGCTCTCTGTGAGAGCTTTCATCATGAGTGGCCCATGGCTGGCATGACATCATG 540
DB 481 ACGGCGCTCTCTGTGAGAGCTTTCATCATGAGTGGCCCATGGCTGGCATGACATCATG 540
QY 541 TTTCCAGCTGGAGTCTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
DB 541 TTTCCAGCTGGAGTCTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 601 TGGAGCCTGAGCGGAGGAGCAGCTGGCCAGACAGCTGCGATGAAGAAGGAGGACCCGG 660
DB 601 TGGAGCCTGAGCGGAGGAGCAGCTGGCCAGACAGCTGCGATGAAGAAGGAGGACCCGG 660
QY 661 TTTCCATCATGCTGGTGGCAATTTGTTTCATCATGCTACCTGCCAGCGTGTCTGCTAGA 720
DB 661 TTTCCATCATGCTGGTGGCAATTTGTTTCATCATGCTACCTGCCAGCGTGTCTGCTAGA 720
QY 721 CTCTATTTCTCTGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 721 CTCTATTTCTCTGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 781 CACATAACCCCTCAGCTTTCACCTACATGAACAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 781 CACATAACCCCTCAGCTTTCACCTACATGAACAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 841 TCAAGCCCTCTCTTCCCAAAATTTACACAAAGCTCAAAATCTGAGTCTGAAACCCAAAG 900
DB 841 TCAAGCCCTCTCTTCCCAAAATTTACACAAAGCTCAAAATCTGAGTCTGAAACCCAAAG 900
QY 901 CAGCAGGAGCACTCAAAACACAAAGGCGGAGAGATGCCAAATTTTCGAACCTCGGTGCG 960
DB 901 CAGCAGGAGCACTCAAAACACAAAGGCGGAGAGATGCCAAATTTTCGAACCTCGGTGCG 960
QY 961 AGGAGTTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 961 AGGAGTTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 1021 CACATTTGTCAGTGGCACTGA 1041
DB 1021 CACATTTGTCAGTGGCACTGA 1041

RESULT 3
AAH51008
ID AAH51008 standard; DNA; 1041 BP.
XX
AC AAH51008;
XX
DT 28-AUG-2001 (first entry)
XX
DE Human nPCR11 coding sequence #2.
XX
KW G protein-coupled receptor; nPCR; seven transmembrane receptor;
KW signal transduction; schizophrenia; thyroid disorder; renal failure;
KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
KW cardiovascular disease; proliferative disorder; hormonal disorder;
KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
KW attention deficit-hyperactivity disorder/attention deficit disorder;
KW Parkinson's disease; migraine; senile dementia; inflammatory disease;
KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;
KW neuroprotective; ds.
XX
OS Homo sapiens.
XX
PN WO200136473-A2.
XX
XX 25-MAY-2001.
XX
PF 16-NOV-2000; 2000WO-US31581.
XX

```
PR 16-NOV-1999; 9905-0165838.
PR 17-NOV-1999; 9905-0166071.
PR 19-NOV-1999; 9905-0166678.
PR 28-DEC-1999; 9905-0173396.
PR 22-FEB-2000; 2000US-0184129.
PR 28-FEB-2000; 2000US-0185421.
PR 28-FEB-2000; 2000US-0185554.
PR 02-MAR-2000; 2000US-0186530.
PR 03-MAR-2000; 2000US-0186811.
PR 09-MAR-2000; 2000US-0188114.
PR 17-MAR-2000; 2000US-0190310.
PR 21-MAR-2000; 2000US-0190800.
PR 20-APR-2000; 2000US-0198568.
PR 02-MAY-2000; 2000US-0201190.
PR 08-MAY-2000; 2000US-0203111.
PR 25-MAY-2000; 2000US-0207094.
XX
XX (PHAA ) PHARMACIA & UPJOHN CO.
XX
XX Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;
PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;
XX
XX WPI; 2001-389826/41.
XX P-PSDB; AAG80968.
XX
XX New G protein-coupled receptor (ngPCR-x) and its encoding
PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -
PT Claim 4; Page 89; 261pp; English.
XX
XX The present invention relates to novel G protein-coupled receptors
CC (ngPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
CC 28, 31-38, 40, 41, 53-60) and their coding sequences. The present
CC sequence is the coding sequence for one such G protein-coupled receptor.
CC GPCRx are also known as seven transmembrane receptors and function in
CC signal transduction. The ngPCRx coding sequences are useful for
CC screening a human to diagnose a disorder affecting the brain or a genetic
CC predisposition, specifically schizophrenia. ngPCRx are useful for
CC identifying compounds useful for treating schizophrenia. Detection of
CC ngPCRx in a sample is useful as a diagnostic tool for diseases or
CC disorders e.g. thyroid disorders, renal failure, rheumatoid arthritis,
CC CNS disorders, infections such as HIV-1, metabolic and cardiovascular
CC diseases, proliferative disorders and hormonal disorders. Modulators of
CC ngPCRx activity have the utility for treating neurological disorders,
CC including schizophrenia, ADHD/ADD (attention deficit-hyperactivity
CC disorder/attention deficit disorder), and neuronal disorders such as
CC Alzheimer's disease, Parkinson's disease, migraine and senile dementia.
CC Additional disorders include inflammatory conditions (e.g. Crohn's
CC disease), rheumatoid arthritis, autoimmune disorders, cancers,
CC respiratory ailments such as asthma, and inflammatory diseases e.g.
CC inflammatory bowel disease.
XX
XX Sequence 1041 BP; 208 A; 294 C; 278 G; 261 T; 0 other;
SQ
Query Match 100.0%; Score 1041; DB 22; Length 1041;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTACACGGGTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATGCCCGCG 60
DB 1 ATGTACACGGGTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATGCCCGCG 60
QY 61 CTGCTCATTTGTGGCTTTGTGCTGGGCGCACCTAGGCAATGGGTCGCTGTGGTTTC 120
DB 61 CTGCTCATTTGTGGCTTTGTGCTGGGCGCACCTAGGCAATGGGTCGCTGTGGTTTC 120
QY 121 TGCTTCACATGAAGACCTGGAAGCCGACGACCTGTTTACCTTTTCAATTTGGCGGTGGCT 180
DB 121 TGCTTCACATGAAGACCTGGAAGCCGACGACCTGTTTACCTTTTCAATTTGGCGGTGGCT 180
QY 181 GATTTCCTCTTATGATCTGCTTCCGTCGCTTTTCCGACACATATTACCTTCAGACGTAGACAC 240
DB 181 GATTTCCTCTTATGATCTGCTTCCGTCGCTTTTCCGACACATATTACCTTCAGACGTAGACAC 240
```

```
QY 241 TGGGCTTTTGGGACATTCCTTCCCGAGTGGGGCTCTTTCACGTTGGCCATGAACAGGGCC 300
DB 241 TGGGCTTTTGGGACATTCCTTCCCGAGTGGGGCTCTTTCACGTTGGCCATGAACAGGGCC 300
QY 301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGGGACAGGTATTTCAAAGTGGTCCACCCC 360
DB 301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGGGACAGGTATTTCAAAGTGGTCCACCCC 360
QY 361 CACACACGGGTGAACACTATCTCCACCCGGGTGGCGCTGGCATCGTCTCACCCCTGTGG 420
DB 361 CACACACGGGTGAACACTATCTCCACCCGGGTGGCGCTGGCATCGTCTCACCCCTGTGG 420
QY 421 GCCTTGGTTCATCTCGGAAACAGTGTATCTTTTGGTGGAGAACCATCTCTCCGTGCAAGAG 480
DB 421 GCCTTGGTTCATCTCGGAAACAGTGTATCTTTTGGTGGAGAACCATCTCTCCGTGCAAGAG 480
QY 481 ACGGCGCTCTCTGTGAGAGCTTCATCATGAGTGGCCCAATGGCTGGCATGACATCATG 540
DB 481 ACGGCGCTCTCTGTGAGAGCTTCATCATGAGTGGCCCAATGGCTGGCATGACATCATG 540
QY 541 TTCAGCTGGAGTCTTTATGCCCCCTCGGCATCATCTTATTTTGGTCTCTTCAAGATTGTT 600
DB 541 TTCAGCTGGAGTCTTTATGCCCCCTCGGCATCATCTTATTTTGGTCTCTTCAAGATTGTT 600
QY 601 TGGAGCCTGAGGGGGAGGAGCAGCTGGCCAGACAGCTCGGATGAAGAGGACCCCG 660
DB 601 TGGAGCCTGAGGGGGAGGAGCAGCTGGCCAGACAGCTCGGATGAAGAGGACCCCG 660
QY 661 TTCATCATGTGTGGCAATTTGTTTTCATCATGCTACCTGCCAGCGTCTGTGCTAGA 720
DB 661 TTCATCATGTGTGGCAATTTGTTTTCATCATGCTACCTGCCAGCGTCTGTGCTAGA 720
QY 721 CTCTATTTCCTCTGGAGGGTGCCTCGAGTGCCTCGCATCCCTCTGTCCATGGGGCCCTG 780
DB 721 CTCTATTTCCTCTGGAGGGTGCCTCGAGTGCCTCGCATCCCTCTGTCCATGGGGCCCTG 780
QY 781 CACATAACCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTCGTGTGTTATTTT 840
DB 781 CACATAACCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTCGTGTGTTATTTT 840
QY 841 TCAGCCCTCTCTTCCCAATTTCTACAACAGCTCAAAATCTGCACTCTGAAACCCCAAG 900
DB 841 TCAGCCCTCTCTTCCCAATTTCTACAACAGCTCAAAATCTGCACTCTGAAACCCCAAG 900
QY 901 CAGCCAGGACACTCAAAACACAAAGCCGAGAGATGCCAATTCGAACTCGGTGCG 960
DB 901 CAGCCAGGACACTCAAAACACAAAGCCGAGAGATGCCAATTCGAACTCGGTGCG 960
QY 961 AGGAGTTGCATCAGTGTGGCAAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020
DB 961 AGGAGTTGCATCAGTGTGGCAAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020
QY 1021 CACATTGTTGAGTGGCACTGA 1041
DB 1021 CACATTGTTGAGTGGCACTGA 1041
RESULT 4
ABR81529
ID ABR81529 standard; DNA; 1050 BP.
XX
AC ABR81529;
XX
XX 28-JAN-2002 (first entry)
XX
XX Human GPCR1a polynucleotide SEQ ID NO 1.
DE
KW Human; GPCR; G-coupled protein-receptor; cardiant; antiarteriosclerotic;
KW anabolic; cytostatic; antiviral; gene therapy; cardiomyopathy; obesity;
KW anorexia; diabetes; osteoporosis; Crohn's disease; multiple sclerosis;
KW asthma; Alzheimer's disease; Parkinson's disorder; Huntington's disease;
KW infection; human immunodeficiency virus; HIV; ds.
```

XX Homo sapiens.
OS W0200174904-A2.
PN 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US10241.
XX 31-MAR-2000; 2000US-193664P.
PR 05-APR-2000; 2000US-194614P.
PR 06-APR-2000; 2000US-195063P.
PR 06-APR-2000; 2000US-195066P.
PR 06-APR-2000; 2000US-195067P.
PR 06-APR-2000; 2000US-195068P.
PR 06-APR-2000; 2000US-195069P.
PR 06-APR-2000; 2000US-195070P.
PR 06-APR-2000; 2000US-195510P.
PR 21-JUL-2000; 2000US-219855P.
PR 27-JUL-2000; 2000US-221284P.
PR 28-JUL-2000; 2000US-221325P.
PR 11-AUG-2000; 2000US-224588P.
PR 11-OCT-2000; 2000US-239613P.
PR 18-JAN-2001; 2001US-262508P.
PR 23-JAN-2001; 2001US-263433P.
PR 23-JAN-2001; 2001US-263604P.
PR 30-JAN-2001; 2001US-265161P.
PR 29-MAR-2001; 2001US-0823172.
XX (CURA-) CURAGEN CORP.
PA Majumder K, Vernet CAM, Casman SJ, Wolenc AR, Spaderna SK;
PI Padigaru M, Mishnu VS, Tchernev VT, Spytek KA, Li L;
PI Baumgartner JC, Gusev VT;
XX WPI; 2001-639351/73.
DR P-PSDB; ABB44522.
XX New human G-protein coupled receptor X, GPCR_X, polypeptide useful in
PT treatment or prevention of GPCR_X associated disorders e.g.
PT cardiomyopathy or atherosclerosis, and to screen for antagonists and
PT agonists useful therapeutically -
XX Claim 9; Page 7; 157pp; English.
XX The invention relates to nucleic acid sequences (ABA81529-ABA81552) that
CC encode G-coupled protein-receptor related polypeptides
CC (ABB44522-ABB44543). The isolated polypeptide having a sequence differing
CC by no more than 15 % of amino acid residues from one of 22 amino acid
CC sequences (or mature forms of the sequences), fully defined in the
CC specification and corresponding to human G-protein coupled receptor X
CC (GPCR_X) polypeptides. The polypeptides have potential cardiant,
CC antiarteriosclerotic, anabolic, cytostatic and antiviral activity. The
CC polypeptides can be administered therapeutically, especially using gene
CC therapy and expressing the encoding DNA in vivo, to treat or prevent
CC GPCR_X-associated disorders, especially in humans. For example, they can
CC be used to treat/prevent cardiomyopathy, atherosclerosis, disorders
CC related to signal processing and metabolic pathway modulation (e.g.
CC obesity, anorexia), diabetes, osteoporosis, Crohn's disease, multiple
CC sclerosis, asthma, cancers, neurodegenerative disorders (e.g. Alzheimer's
CC disease, Parkinson's disorder, Huntington's disease), immune disorders,
CC haematopoietic disorders, developmental diseases, neurological disorders,
CC bacterial, fungal, protozoal and viral infections (e.g. with human
CC immunodeficiency virus (HIV)-1 or HIV-2). They can be used diagnostically
CC to determine the presence of or predisposition to a disease associated
CC with altered levels of the polypeptide in mammals (especially humans) by
CC detecting alterations in polypeptide expression levels relative to
CC control samples. They are useful to identify agents binding polypeptide
CC (e.g. cellular receptors or downstream effectors) and/or agents
CC modulating cellular polypeptide expression or activity, useful as
CC antagonists and agonists in disease treatment.
XX Sequence 1050 BP; 211 A; 299 C; 279 G; 261 T; 0 other;

Query Match 100.0%; Score 1041; DB 22; Length 1050;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGTACAACGGGTGCTGCTGCCGATCGAGGGGGACACCATCTCCAGGTGATGCCGCCG	60
DB	6	ATGTACAACGGGTGCTGCTGCCGATCGAGGGGGACACCATCTCCAGGTGATGCCGCCG	65
QY	61	CTGCTCATTTGTCGCTTGGGCGCACTAGGCAATGGGTCGCCCTGTGTTTC	120
DB	66	CTGCTCATTTGTCGCTTGGGCGCACTAGGCAATGGGTCGCCCTGTGTTTC	125
QY	121	TGCTTCCACATGAAGACCTGGAAGCCGACGACCTTTTACCTTTTCAATTTGGCGTGGCT	180
DB	126	TGCTTCCACATGAAGACCTGGAAGCCGACGACCTTTTACCTTTTCAATTTGGCGTGGCT	185
QY	181	GATTTCTCTTATGATCTGCCCTTTTCGGACAGACTATTACCTTCAGACGTAGACAC	240
DB	186	GATTTCTCTTATGATCTGCCCTTTTCGGACAGACTATTACCTTCAGACGTAGACAC	245
QY	241	TGGGCTTTTGGGACATTCCTCGCGAGTGGGCTCTTACGTTGGCCATGAACAGGCCC	300
DB	246	TGGGCTTTTGGGACATTCCTCGCGAGTGGGCTCTTACGTTGGCCATGAACAGGCCC	305
QY	301	GGGAGCATCGTTTCCCTTACGGTGGTGGCTCGGACAGGTATTTCAAAGTGGTCCACCCC	360
DB	306	GGGAGCATCGTTTCCCTTACGGTGGTGGCTCGGACAGGTATTTCAAAGTGGTCCACCCC	365
QY	361	CACCACGCGTGAACACTATCTCCACCCGGTGGCGCTGGCATGCTGCACCCCTGTGG	420
DB	366	CACCACGCGTGAACACTATCTCCACCCGGTGGCGCTGGCATGCTGCACCCCTGTGG	425
QY	421	GCCTGGTCACTCTGGGAACAGTGTATCTTTTGTCTGGAGAACCATCTCTCGGTGCAAGAG	480
DB	426	GCCTGGTCACTCTGGGAACAGTGTATCTTTTGTCTGGAGAACCATCTCTCGGTGCAAGAG	485
QY	481	ACGGCGCTCTCTGTGAGAGCTTTCATCATGGAGTGGGCCAATGGCTGGCATGACATCATG	540
DB	486	ACGGCGCTCTCTGTGAGAGCTTTCATCATGGAGTGGGCCAATGGCTGGCATGACATCATG	545
QY	541	TTCCAGCTGGAGTCTTTATGCCCTCGGCATCATCTTATTTGCTCCTTCAAGATTGTT	600
DB	546	TTCCAGCTGGAGTCTTTATGCCCTCGGCATCATCTTATTTGCTCCTTCAAGATTGTT	605
QY	601	TGAGCCTGAGCGGAGGAGCAGCTGGCCAGACAGCTCGGATGAAGAAGCGACCCGG	660
DB	606	TGAGCCTGAGCGGAGGAGCAGCTGGCCAGACAGCTCGGATGAAGAAGCGACCCGG	665
QY	661	TTCATCATGTTGTTGGCAATTGTTTTCATCAGATGCTACCTGCCAGCGTGTCTGTAGTA	720
DB	666	TTCATCATGTTGTTGGCAATTGTTTTCATCAGATGCTACCTGCCAGCGTGTCTGTAGTA	725
QY	721	CTCTATTTCTCTGGAGCGTGCCTCGAGTGCCTGGATCCCTGTGTCATGGGGCCCTG	780
DB	726	CTCTATTTCTCTGGAGCGTGCCTCGAGTGCCTGGATCCCTGTGTCATGGGGCCCTG	785
QY	781	CACATACCTCTCAGCTTCACCTACATGAACAGATGCTGGATCCCTGTGTTATTTT	840
DB	786	CACATACCTCTCAGCTTCACCTACATGAACAGATGCTGGATCCCTGTGTTATTTT	845
QY	841	TCAAGCCCTCTCTTCCCAAAATCTACAACAGCTCAAAATCTGCAGCTGAAACCCCAAG	900
DB	846	TCAAGCCCTCTCTTCCCAAAATCTACAACAGCTCAAAATCTGCAGCTGAAACCCCAAG	905
QY	901	CAGCCAGGACACTCAAAACACAAAGCCCGGAAGAGATGCCAAATTTCCAAACCTCGGTGCG	960
DB	906	CAGCCAGGACACTCAAAACACAAAGCCCGGAAGAGATGCCAAATTTCCAAACCTCGGTGCG	965
QY	961	AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC	1020
DB	966	AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC	1025

QY 1021 CACATTTGAGTGGCACTGA 1041
DB 1026 CACATTTGAGTGGCACTGA 1046

RESULT 5
ID ABA81530
XX ABA81530 standard; DNA; 1050 BP.

AC ABA81530;
XX
DT 28-JAN-2002 (first entry)
XX
DE Human GPCR1b polynucleotide SEQ ID NO 3.

XX Human; GPCR; G-coupled protein-receptor; cardiant; antiarteriosclerotic;
KW anabolic; cytosolic; antiviral; gene therapy; cardiomyopathy; obesity;
KW anorexia; diabetes; osteoporosis; Crohn's disease; multiple sclerosis;
KW asthma; Alzheimer's disease; Parkinson's disorder; Huntington's disease;
KW Infection; human immunodeficiency virus; HIV; ds.

XX Homo sapiens.
OS
XX
XX W0200174904-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US10241.

XX 31-MAR-2000; 2000US-193664P.
PR 05-APR-2000; 2000US-194614P.
PR 06-APR-2000; 2000US-195063P.
PR 06-APR-2000; 2000US-195066P.
PR 06-APR-2000; 2000US-195067P.
PR 06-APR-2000; 2000US-195068P.
PR 06-APR-2000; 2000US-195069P.
PR 06-APR-2000; 2000US-195070P.
PR 06-APR-2000; 2000US-195510P.
PR 21-JUL-2000; 2000US-219855P.
PR 27-JUL-2000; 2000US-221284P.
PR 28-JUL-2000; 2000US-221325P.
PR 11-AUG-2000; 2000US-224588P.
PR 11-OCT-2000; 2000US-239613P.
PR 18-JAN-2001; 2001US-262508P.
PR 23-JAN-2001; 2001US-263433P.
PR 23-JAN-2001; 2001US-263604P.
PR 30-JAN-2001; 2001US-265161P.
PR 29-MAR-2001; 2001US-0823172.

(CURA-) CURAGEN CORP.

XX
XX Majumder K, Vernet CAM, Casman SJ, Wolenc AR, Spaderna SK;
PI Padigaru M, Mishnu VS, Tchernev VT, Spytek KA, Li L;
PI Baumgartner JC, Gusev VY;

XX WPI; 2001-639351/73.
XX P-PSDB; ABB44522.

XX New human G-protein coupled receptor X, GPCR, polypeptide useful in
PT treatment or prevention of GPCR associated disorders e.g.
PT cardiomyopathy or atherosclerosis, and to screen for antagonists and
PT agonists useful therapeutically

XX Claim 9; Page 10; 157pp; English.

XX The invention relates to nucleic acid sequences (ABA81529-ABA81552) that
CC encode G-coupled protein-receptor related polypeptides
CC (ABB44522-ABB44543). The isolated polypeptide having a sequence differing
CC by no more than 15 % of amino acid residues from one of 22 amino acid
CC sequences (or mature forms of the sequences), fully defined in the
CC specification and corresponding to human G-protein coupled receptor X
CC (GPCR) polypeptides. The polypeptides have potential cardiant,
XX antiarteriosclerotic, anabolic, cytostatic and antiviral activity. The

CC polypeptides can be administered therapeutically, especially using gene
CC therapy and expressing the encoding DNA in vivo, to treat or prevent
CC GPCR-associated disorders, especially in humans. For example, they can
CC be used to treat/prevent cardiomyopathy, atherosclerosis, disorders
CC related to signal processing and metabolic pathway modulation (e.g.
CC obesity, anorexia), diabetes, osteoporosis, Crohn's disease, multiple
CC sclerosis, asthma, cancers, neurodegenerative disorders (e.g. Alzheimer's
CC disease, Parkinson's disorder, Huntington's disease), immune disorders,
CC haematopoietic disorders, developmental diseases, neurological disorders,
CC bacterial, fungal, protozoal and viral infections (e.g. with human
CC immunodeficiency virus (HIV)-1 or HIV-2). They can be used diagnostically
CC to determine the presence of or predisposition to a disease associated
CC with altered levels of the polypeptide in mammals (especially humans) by
CC detecting alterations in polypeptide expression levels relative to
CC control samples. They are useful to identify agents binding polypeptide
CC (e.g. cellular receptors or downstream effectors) and/or agents
CC modulating cellular polypeptide expression or activity, useful as
CC antagonists and agonists in disease treatment.

XX
SQ Sequence 1050 BP; 211 A; 298 C; 279 G; 262 T; 0 other;

Query Match 100.0%; Score 1041; DB 22; Length 1050;
Best Local Similarity 100.0%; Pred No. 0;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTACAACGGGTGCTGCTGCCGCATCGAGGGGACACCATCTCCAGAGTGATCCGCCG 60
DB 6 ATGTACAACGGGTGCTGCTGCCGCATCGAGGGGACACCATCTCCAGAGTGATCCGCCG 65
QY 61 CTGCTCATTTGGGCTTTTGGCTGGGGCCACTAGGCAATGGGGTCGCCCTGTGGTTTC 120
DB 66 CTGCTCATTTGGGCTTTTGGCTGGGGCCACTAGGCAATGGGGTCGCCCTGTGGTTTC 125
QY 121 TGCTTCCACATGAAGACCTGGAAGCCAGCAGCTGTTTACCTTTTCAATTTGGCGTGGCT 180
DB 126 TGCTTCCACATGAAGACCTGGAAGCCAGCAGCTGTTTACCTTTTCAATTTGGCGTGGCT 185
QY 181 GATTTCTCTTATGATCTGCCTGCTGCTGCGGACAGACTATTACCTCAGACGTAGACAC 240
DB 186 GATTTCTCTTATGATCTGCCTGCTGCTGCGGACAGACTATTACCTCAGACGTAGACAC 245
QY 241 TGGGCTTTTGGGGACATTCCTGCCGAGTGGGGCTCTTCAGCTTGGCCATGAACAGGCC 300
DB 246 TGGGCTTTTGGGGACATTCCTGCCGAGTGGGGCTCTTCAGCTTGGCCATGAACAGGCC 305
QY 301 GGGAGCATCGTGTCTTACGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC 360
DB 306 GGGAGCATCGTGTCTTACGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC 365
QY 361 CACCACCGGTGAACACTATCTCCACCGGGTGGCGGTGGCATCGTGTGCACCCCTGTGG 420
DB 366 CACCACCGGTGAACACTATCTCCACCGGGTGGCGGTGGCATCGTGTGCACCCCTGTGG 425
QY 421 GCCCTGGTCACTCTGGGAACAGTGTATCTTTTGTGGAGAACCATCTCTGGGTGCAAGAG 480
DB 426 GCCCTGGTCACTCTGGGAACAGTGTATCTTTTGTGGAGAACCATCTCTGGGTGCAAGAG 485
QY 481 ACGCCGCTCTCTGTGAGAGCTTTCATGAGTGGCGCAATGGTGGCATGACATCATG 540
DB 486 ACGCCGCTCTCTGTGAGAGCTTTCATGAGTGGCGCAATGGTGGCATGACATCATG 545
QY 541 TTCCAGCTGGAGTCTTTTATGCCCTCGGCATCATCTTATTTTGTCTCTCAAGATTGTT 600
DB 546 TTCCAGCTGGAGTCTTTTATGCCCTCGGCATCATCTTATTTTGTCTCTCAAGATTGTT 605
QY 601 TGGAGCTGAGGGGAGGACAGCAGTGGCCAGACAGGCTCGGATGAAGAGGCCGCCCG 660
DB 606 TGGAGCTGAGGGGAGGACAGCAGTGGCCAGACAGGCTCGGATGAAGAGGCCGCCCG 665
QY 661 TTTCATCATGGTGGGAATTTGTTCATCATCATGCTACCTGCCAGCGTCTCTGCTAGA 720
DB 666 TTTCATCATGGTGGGAATTTGTTCATCATCATGCTACCTGCCAGCGTCTCTGCTAGA 725


```
QY 421 GCCCTGTCATCTCTGGGACAGTGATCTTTTGTGAGAACCATCTCTCGTGCAGAGAG 480
|
|
|
Db 463 GCCCTGTCATCTCTGGGACAGTGATCTTTTGTGAGAACCATCTCTCGTGCAGAGAG 522
|
|
|
QY 481 ACGCCCTCTCTCTGTGAGAGCTTCATCATGAGTGGCGCCAAATGGCTGCATGACATCATG 540
|
|
|
Db 523 ACGCCCTCTCTCTGTGAGAGCTTCATCATGAGTGGCGCCAAATGGCTGCATGACATCATG 582
|
|
|
QY 541 TTCAGCTGGAGTCTTTATGCCCCCTCGGCATCATCTATTTTGTCTCTCAAGATTGTT 600
|
|
|
Db 583 TTCAGCTGGAGTCTTTATGCCCCCTCGGCATCATCTATTTTGTCTCTCAAGATTGTT 642
|
|
|
QY 601 TGGAGCTGAGGGGAGGCGAGCAGCTGCCAGACAGAGCTCGGATGAAGAGGGGACCCGG 660
|
|
|
Db 643 TGGAGCTGAGGGGAGGCGAGCAGCTGCCAGACAGAGCTCGGATGAAGAGGGGACCCGG 702
|
|
|
QY 661 TTCATCATGTTGGTGGCAATGTTTTCATCATGCTGACCTGACCGCGTGTGCTGCTGAGA 720
|
|
|
Db 703 TTCATCATGTTGGTGGCAATGTTTTCATCATGCTGACCTGACCGCGTGTGCTGCTGAGA 762
|
|
|
QY 721 CTCTATTTCCTCTGGAGCGTGCCTCGAGTGCCTGCGATCCCTCTGTCCATGGGGCCCTG 780
|
|
|
Db 763 CTCTATTTCCTCTGGAGCGTGCCTCGAGTGCCTGCGATCCCTCTGTCCATGGGGCCCTG 822
|
|
|
QY 781 CACATAACCTTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTCTGTGTTATTTT 840
|
|
|
Db 823 CACATAACCTTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTCTGTGTTATTTT 882
|
|
|
QY 841 TCAAGCCCCCTCTTCCCAATTTTACAACAAGCTCAAAATGTCAGTCTGAAACCCCAAG 900
|
|
|
Db 883 TCAAGCCCCCTCTTCCCAATTTTACAACAAGCTCAAAATGTCAGTCTGAAACCCCAAG 942
|
|
|
QY 901 CAGCCAGGACACTCAAAAACACAAAGCCGGAAGAGATGCCAATTTGCAACCTCGTGC 960
|
|
|
Db 943 CAGCCAGGACACTCAAAAACACAAAGCCGGAAGAGATGCCAATTTGCAACCTCGTGC 1002
|
|
|
QY 961 AGGAGTTCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020
|
|
|
Db 1003 AGGAGTTCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1062
|
|
|
QY 1021 CACATTGTTGAGTGGCACTGA 1041
|
|
|
Db 1063 CACATTGTTGAGTGGCACTGA 1083
|
|
|
RESULT 7
AAS18501
ID AAS18501 standard; cDNA; 1730 BP.
XX
AC AAS18501;
XX
DT 26-FEB-2002 (first entry)
XX
DE cDNA encoding HM74-like G-protein coupled receptor (GPCR).
XX
KW HM74-like GPCR; G-protein coupled receptor; antibacterial; fungicide;
KW protozoacide; analgesic; cytostatic; neuroleptic; nootropic; anorexia;
KW anticonvulsant; tranquilizer; viral infection; pain; cancer; anorexia;
KW bulimia; asthma; central nervous system disease; CNS disease;
KW cardiovascular disease; hypotension; hypertension; angina pectoris;
KW myocardial infarction; urinary retention; osteoporosis; ulcer; asthma;
KW inflammation; allergy; benign prostatic hypertrophy; multiple sclerosis;
KW psychotic disorder; neurological disorder; dyskinesia;
KW Huntington's disease; Tourette's syndrome; anxiety; schizophrenia;
KW manic depression; delirium; dementia; mental retardation; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 464..1504
FT /*tag= a
FT /*product= "HM74-like_GPCR"
```

```
FT
XX
PN
XX
XX
PD 18-OCT-2001.
XX
XX 04-APR-2001; 2001WO-EP03811.
XX
XX 05-APR-2000; 2000US-194701P.
XX
XX (FARB ) BAYER AG.
XX
XX Xiao Y;
XX
XX WPI: 2002-049147/06.
XX
XX P-PSDB; AAU11401.
XX
XX
XX Novel isolated polynucleotide, useful for treating infection, pain,
XX cancer, asthma, hypertension, myocardial infarction, urinary retention,
XX osteoporosis, encodes the human HM74-like G-protein coupled receptor
XX polypeptide
XX
XX Claim 1; Fig 1; 77pp; English.
XX
XX The invention describes a novel isolated polynucleotide (I) encoding a
XX human HM74-like G-protein coupled receptor (GPCR) polypeptide. Reagents
XX that regulate HM74-like GPCR are useful for modulating the activity of
XX the protein in a disease selected from bacterial, fungal, protozoan, and
XX viral infection, pain, cancer, anorexia, bulimia, asthma, central nervous
XX system (CNS) disease, cardiovascular disease, hypotension, hypertension,
XX angina pectoris, myocardial infarction, urinary retention, osteoporosis,
XX ulcer, asthma, inflammation, allergy, benign prostatic hypertrophy,
XX multiple sclerosis and dyskinesia such as Huntington's disease and
XX Tourette's syndrome. The composition is also useful for treating
XX psychotic and neurological disorders such as anxiety, schizophrenia,
XX manic depression, delirium, dementia and severe mental retardation. (I)
XX or the HM74-like GPCR polypeptide are also useful for treating the above
XX mentioned diseases. (I) is useful in a diagnostic assay for detecting
XX diseases, susceptibility to diseases and abnormalities related to the
XX presence of mutations in the nucleic acid sequences which encode a GPCR.
XX The polypeptide is useful to identify test compounds which may act as
XX agonists or antagonists at the receptor site and which can be regulated
XX to provide therapeutic effects. The polypeptide is also useful as a bait
XX protein in a two-hybrid or three-hybrid assay, and to immunise a mammal
XX for production of polyclonal antibodies. This sequence encodes the human
XX HM74-like GPCR described in the method of the invention.
XX
XX Sequence 1730 BP; 361 A; 494 C; 444 G; 431 T; 0 other;
XX
XX Query Match 100.0%; Score 1041; DB 24; Length 1730;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTACAAACGGGTGCTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 60
Db 464 ATGTACAAACGGGTGCTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 523
QY 61 CTGCTCATTTGCGCTTGTGCTGGCGCACTAGGCAATGGGTCGCCCTGTGTGTTTC 120
Db 524 CTGCTCATTTGCGCTTGTGCTGGCGCACTAGGCAATGGGTCGCCCTGTGTGTTTC 583
QY 121 TGGTTCACATGAAGACCTGGAAGCCCGACGACTGTTTACCTTTTCAATTTGGCGTGGCT 180
Db 584 TGGTTCACATGAAGACCTGGAAGCCCGACGACTGTTTACCTTTTCAATTTGGCGTGGCT 643
QY 181 GATTTCCTCTTATGATCTGCCTGCGCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
Db 644 GATTTCCTCTTATGATCTGCCTGCGCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 703
QY 241 TGGGCTTTTGGGACATTTCCCTGCCGAGTGGGGCTCTTACGTTGGCCATGACAGAGCC 300
Db 704 TGGGCTTTTGGGACATTTCCCTGCCGAGTGGGGCTCTTACGTTGGCCATGACAGAGCC 763
```

/note= "G-protein coupled receptor"

QY 301 GGGAGCATGTTCTTACGGTGTGGTGGCGACAGGATATTTCAAAGTGTCCACCCC 360
DB 764 GGGAGCATGTTCTTACGGTGTGGTGGCGACAGGATATTTCAAAGTGTCCACCCC 823
QY 361 CACCAGCGGTGAACACTATCTCCACCGGTGGCGGTGGCATCTGTGCACCCCTGTGG 420
DB 824 CACCAGCGGTGAACACTATCTCCACCGGTGGCGGTGGCATCTGTGCACCCCTGTGG 883
QY 421 GCCCTGTCATCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAGAG 480
DB 884 GCCCTGTCATCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAGAG 943
QY 481 ACGCGCGTCTCTGTGAGAGCTTCATCATGGAGTGGCCCAATGGCTGGCATGACATCATG 540
DB 944 ACGCGCGTCTCTGTGAGAGCTTCATCATGGAGTGGCCCAATGGCTGGCATGACATCATG 1003
QY 541 TTCCAGCTGGAGTCTTTATGCCCCCTCGGCATCATCTTATTTTGGCTCTCAAGATTGTT 600
DB 1004 TTCCAGCTGGAGTCTTTATGCCCCCTCGGCATCATCTTATTTTGGCTCTCAAGATTGTT 1063
QY 601 TGGAGCTGAGCGGAGGAGCTGGCGACAGAGCTCGGATGAAGAAGCGACCCCG 660
DB 1064 TGGAGCTGAGCGGAGGAGCTGGCGACAGAGCTCGGATGAAGAAGCGACCCCG 1123
QY 661 TTCATCATGGTGGTGCATATGTTTCATCATCATGCTACCTGCCAGCGTGTGTGCTAGA 720
DB 1124 TTCATCATGGTGGTGCATATGTTTCATCATCATGCTACCTGCCAGCGTGTGTGCTAGA 1183
QY 721 CTCATTTCTCTGGAGCGTGGCTCGAGTGGCTGGATGCCCTCTCTCCATGGGCCCCTG 780
DB 1184 CTCATTTCTCTGGAGCGTGGCTCGAGTGGCTGGATGCCCTCTCTCCATGGGCCCCTG 1243
QY 781 CACATAACCTCAGCTTACCTACATGAACAGCATCTGGATCCCTCGTGTATTATTTT 840
DB 1244 CACATAACCTCAGCTTACCTACATGAACAGCATCTGGATCCCTCGTGTATTATTTT 1303
QY 841 TCAAGCCCTCTCTTCCCAATTTCTACAAAGCTCAAAATCTGAGTCTGAAACCCAAAG 900
DB 1304 TCAAGCCCTCTCTTCCCAATTTCTACAAAGCTCAAAATCTGAGTCTGAAACCCAAAG 1363
QY 901 CAGCAGGAGACTCAAAACACAAAGCGGAGAGATGCCAATTTTCAACCTCGGTCCG 960
DB 1364 CAGCAGGAGACTCAAAACACAAAGCGGAGAGATGCCAATTTTCAACCTCGGTCCG 1423
QY 961 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCAGTCTGATGGCAATGGGATCCC 1020
DB 1424 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCAGTCTGATGGCAATGGGATCCC 1483
QY 1021 CACATTGTTGAGTGGCACTGA 1041
DB 1484 CACATTGTTGAGTGGCACTGA 1504

RESULT 8
AAS12582
ID AAS12582 standard; DNA; 2331 BP.
XX
AC AAS12582;
XX
DT 19-DEC-2001 (first entry)
XX
DE Gene encoding novel human G protein-coupled receptor (GPCR).
XX
KW Human: G-protein coupled receptor; GPCR; chemokine receptor; protease;
KW hyperproliferative disorder; neurological disorder; psychiatric disease;
KW inflammatory disorder; respiratory disorder; gene therapy; ds.
OS Homo sapiens.
XX
XX W0200173029-A2.
XX
PD 04-OCT-2001.
XX

PF 27-MAR-2001; 2001WO-US09522.
XX
PR 27-MAR-2000; 2000US-192419P.
PR 06-SEP-2000; 2000US-230459P.
XX 20-SEP-2000; 2000US-0666535.
PA (PEKE) PE CORP NY.
XX
PI Ye J, Cravchik A, Di Francesco V, Beasley EM;
XX WPI; 2001-616503/71.
DR
XX
PT Novel human G-protein coupled receptor proteins and nucleic acid
PT molecules encoding the protein for use in developing human therapeutics
PT and diagnostic compositions and for identifying modulators of the
PT protein
XX
PS Claim 23; Fig 3; 66pp; English.
XX
CC The present invention relates to the isolation of a novel human G-protein
CC coupled receptor (GPCR) which is related to the chemokine receptor
CC subfamily. The cDNA and gene sequences encoding for GPCR are also
CC given in the invention. The sequences of the invention are useful
CC for diagnosing and treating diseases or conditions mediated by human
CC proteases. Such diseases include hyperproliferative disorders
CC (e.g. hyperplasia), neurological disorders (e.g. Parkinson's disease),
CC psychiatric diseases (e.g. schizophrenia), inflammatory disorders
CC (e.g. diabetes) and respiratory disorders (e.g. adult respiratory
CC distress syndrome, ARDS). The GPCR protein is also useful for identifying
CC a modulator of the expression of the protein. It also serves as a target
CC for identifying agents for use in mammalian therapeutic applications,
CC e.g. a human drug, particularly modulating a biological or pathological
CC response in a cell or tissue that expresses the protein, in biological
CC assays related to GPCRs that are related to members of the chemokine
CC receptor subfamily, in drug screening assays and in competition binding
CC assays. GPCR is also useful in diagnosing a disease or predisposition to
CC a disease mediated by the peptide, in pharmacogenomic analysis. The
CC polynucleotide sequences can also be used in gene therapy. The present
CC sequence represents the human GPCR gene sequence of the invention.
XX
SQ Sequence 2331 BP; 497 A; 620 C; 592 G; 622 T; 0 other;
Query Match 100.0%; Score 1041; DB 22; Length 2331;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTACAAGGGTCTGCTGCCCATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 60
DB 311 ATGTACAAGGGTCTGCTGCCCATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 370
QY 61 CTGCTCATTTGGCCCTTTGTGCTGGCGCACTAGGCAATGGGCTCGCCCTGTGTGTTTC 120
DB 371 CTGCTCATTTGGCCCTTTGTGCTGGCGCACTAGGCAATGGGCTCGCCCTGTGTGTTTC 430
QY 121 TGCTTCCACATGAAGACCTTGAAGCCAGACACTGTTTACCTTTTCAATTTGGCCGTGGCT 180
DB 431 TGCTTCCACATGAAGACCTTGAAGCCAGACACTGTTTACCTTTTCAATTTGGCCGTGGCT 490
QY 181 GATTTCTCTTATGATCTGCCCTTTTCGGACAGACTATTACCTCAGACATGACAC 240
DB 491 GATTTCTCTTATGATCTGCCCTTTTCGGACAGACTATTACCTCAGACATGACAC 550
QY 241 TGGGCTTTTGGGACATCTCCCTCGGAGTGGGCTCTTACAGTTGGCATGAACAGGCC 300
DB 551 TGGGCTTTTGGGACATCTCCCTCGGAGTGGGCTCTTACAGTTGGCATGAACAGGCC 610
QY 301 GGGAGCATCGTGTCTTACGGTGGTGGCTGCGGACAGGATTTTCAAAGTGTGTCACCCC 360
DB 611 GGGAGCATCGTGTCTTACGGTGGTGGCTGCGGACAGGATTTTCAAAGTGTGTCACCCC 670
QY 361 CACCAGCGGTGAACACTATCTCCACCGGTGGCGGTGGCATCTGTGCACCCCTGTGG 420
DB 671 CACCAGCGGTGAACACTATCTCCACCGGTGGCGGTGGCATCTGTGCACCCCTGTGG 730

DT	XX	23-OCT-2001	(first entry)
DE	XX	Human cDNA encoding G-protein coupled receptor, hRUP25.	
XX	XX	Human; G-protein coupled receptor; GPCR; hRUP25; agonist;	
KW	XX	Inverse agonist; lung cancer; ss.	
OS	XX	Homo sapiens.	
PH	FT	Key Location/Qualifiers	
FT	FT	CDS 1..1092	
FT	FT	/*tag= a	
XX	XX	/product= "hRUP25"	
PN	PN	WO200136471-A2.	
XX	PD	25-MAY-2001.	
PF	XX	16-NOV-2000; 2000WO-US31509.	
PR	PR	17-NOV-1999; 99US-0166088.	
PR	PR	17-NOV-1999; 99US-0166099.	
PR	PR	17-NOV-1999; 99US-0166369.	
PR	PR	23-DEC-1999; 99US-0171900.	
PR	PR	23-DEC-1999; 99US-0171901.	
PR	PR	23-DEC-1999; 99US-0171902.	
PR	PR	11-FEB-2000; 2000US-0181749.	
PR	PR	14-MAR-2000; 2000US-0189258.	
PR	PR	14-MAR-2000; 2000US-0189259.	
PR	PR	10-APR-2000; 2000US-0195898.	
PR	PR	10-APR-2000; 2000US-0195899.	
PR	PR	10-APR-2000; 2000US-0196078.	
PR	PR	28-APR-2000; 2000US-0200419.	
PR	PR	12-MAY-2000; 2000US-0203630.	
PR	PR	12-JUN-2000; 2000US-0210741.	
PR	PR	12-JUN-2000; 2000US-0210982.	
PR	PR	21-AUG-2000; 2000US-0226760.	
PR	PR	26-SEP-2000; 2000US-0235418.	
PR	PR	26-SEP-2000; 2000US-0235779.	
PR	PR	20-OCT-2000; 2000US-0242332.	
PR	PR	20-OCT-2000; 2000US-0242343.	
PA	PA	(AREN-) ARENA PHARM INC.	
PI	PI	Chen R, Dang HT, Lowitz KP;	
XX	XX	WPI: 2001-355616/37.	
DR	DR	P-PFDB: AAU04379.	
XX	PT	Endogenous and non-endogenous versions of human G-protein coupled receptors for direct identification of candidate compounds as agonists, inverse agonists or partial agonists for use as therapeutic agents -	
PS	PS	Claim 71; Page 120-121; 159pp; English.	
CC	CC	The sequence encodes a human G-protein coupled receptor (GPCR), hRUP25. The endogenous and non-endogenous, constitutively activated versions of human G-protein coupled receptors (GPCR), are useful for direct identification of candidate compounds as receptor agonists, inverse agonists or partial agonists having applicability as therapeutic agents for treating diseases related to GPCR, e.g. lung cancer.	
CC	CC	Non-endogenous version of human GPCRs are also utilized in research settings and in vitro and in vivo system, incorporating GPCRs can be utilised to elucidate and understand the roles these receptors play in the human condition, both normal and diseased.	
XX	SQ	Sequence 1092 BP; 223 A; 320 C; 273 G; 276 T; 0 other;	
Query Match		3.1%; Score 32; DB 22; Length 1092;	
Best Local Similarity		100.0%; Pred. No. 8.9e-06;	
Mismatches		Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	61	CTGCTCATTTGGCCTTGTGCTGGGGGCACACTAGGAATGGGTCGCCCTGTGTGTTTC 120	
Dd	1	CTGCTCATTTGGCCTTGTGCTGGGGGCACACTAGGAATGGGTCGCCCTGTGTGTTTC 60	
QY	121	TGCTTCCACATGAAGACCTCGAAGCCCAGCACGTGTTTACCTTTCAATTGGCCGTGGCT 180	
Dd	61	TGCTTCCACATGAAGACCTCGAAGCCCAGCACGTGTTTACCTTTCAATTGGCCGTGGCT 120	
QY	181	GATTTCCTCCTTATGATCTGCCCTGCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240	
Dd	121	GATTTCCTCCTTATGATCTGCCCTGCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 180	
QY	241	TGGCTTTTGGGACATTCCTGCCAGTGCGGCTCTTACGCTGGGCCATGAACAGGGCC 300	
Dd	181	TGGCTTTTGGGACATTCCTGCCAGTGCGGCTCTTACGCTGGGCCATGAACAGGGCC 240	
QY	301	GGAGACATCGTGTTCCTTAGCGGTGGTGGTGGCGGCTGGCATCTCTGCAACGTCACACC 360	
Dd	241	GGAGACATCGTGTTCCTTAGCGGTGGTGGTGGCGGCTGGCATCTCTGCAACGTCACACC 300	
QY	361	CACCACGCTGAACACTATCTCCACCCGGTGGCGGCTGGCATCTCTGCAACGTCACACC 420	
Dd	301	CACCACGCTGAACACTATCTCCACCCGGTGGCGGCTGGCATCTCTGCAACGTCACACC 360	
QY	421	GCCTGTGTCATCTGGAAACAGTGTATCTTTTGGTGGAAACCATCTCTGCGTGCAAGAG 480	
Dd	361	GCCTGTGTCATCTGGAAACAGTGTATCTTTTGGTGGAAACCATCTCTGCGTGCAAGAG 420	
QY	481	ACGGCTGCTCTCTGAGAGCTTCATCTGAGTGGAGTGGCGGCAATGGCTGGCATCATG 540	
Dd	421	ACGGCTGCTCTCTGAGAGCTTCATCTGAGTGGAGTGGCGGCAATGGCTGGCATCATG 480	
QY	541	TTCCAGCTGAGTCTTTTATGCCCCCTCGGCATCATCTTTTTCCTTCAAGATTGTT 600	
Dd	481	TTCCAGCTGAGTCTTTTATGCCCCCTCGGCATCATCTTTTTCCTTCAAGATTGTT 540	
QY	601	TGAGCCTGAGGGGAGGAGCAGCTGGCCAGACAGCTCGGATGAAGAAGCGACCCGG 660	
Dd	541	TGAGCCTGAGGGGAGGAGCAGCTGGCCAGACAGCTCGGATGAAGAAGCGACCCGG 600	
QY	661	TTTCATCATGTGTGGCAATTGTTTCATCATCATGTCTGCTGCCAGCGTGTCTGTAGA 720	
Dd	601	TTTCATCATGTGTGGCAATTGTTTCATCATCATGTCTGCTGCCAGCGTGTCTGTAGA 660	
QY	721	CTCTATTCTCTGGAGCTGCTCGAGTGGCTGGGATCCCTCTCTCCATGGGCGCTG 780	
Dd	661	CTCTATTCTCTGGAGCTGCTCGAGTGGCTGGGATCCCTCTCTCCATGGGCGCTG 720	
QY	781	CACATAACCTCAGCTTACCTACATGACAGCATGCTGGATCCCTGTGTATTATTTT 840	
Dd	721	CACATAACCTCAGCTTACCTACATGACAGCATGCTGGATCCCTGTGTATTATTTT 780	
QY	841	TCAGGCCCTCTTCCCAAATCTACACAAGCTCAAAATCTGCACTGTAACCCCAAG 900	
Dd	781	TCAGGCCCTCTTCCCAAATCTTCAACAAGCTCAAAATCTGCACTGTAACCCCAAG 840	
QY	901	CAGCAGGACACTCAAAAACACAAAGCGGGAAGAGATGCCAATTCG 948	
Dd	841	CAGCAGGACACTCAAAAACACAAAGCGGGAAGAGATGCCAATTCG 888	
RESULT 11			
AAS07952			
ID	AAS07952	standard; cDNA; 1092 BP.	
XX	AC		
XX	AC	AAS07952;	
XX	XX		

Db 838 CTCAGCTTCACCTACATGAACAGCATGCTGGA 869
|||||

RESULT 12

AAA30658
ID AAA30658 standard; cDNA; 1164 BP.

AC AAA30658;

DT 21-AUG-2000 (first entry)

DE Human G protein-coupled receptor HM74 cDNA.

XX G protein-coupled receptor; GPCR; constitutively active;

KW intracellular loop 3; transmembrane domain 6; drug screening;

KW agonist; antagonist; ss.

XX Homo sapiens.

OS WO200022129-A1.

PN 20-APR-2000.

PD 12-OCT-1999; 99WO-US23938.

XX 13-OCT-1998; 98US-0170496.

XX (AREN-) ARENA PHARM INC.

PA Behan DP, Chalmers DT, Liaw CW;

XX WPI: 2000-329165/28.

XX P-PSDB; AAY90637.

DR Non-endogenous constitutively activated human G protein-coupled

PT receptors, useful for identifying agonists for use as pharmaceutical

PT agents -

XX Example 1; Page 185; 341pp; English.

XX The invention relates to constitutively active, non-endogenous versions

CC of endogenous human orphan G protein-coupled receptors (GPCRs, AAY90643-

CC AAY90677 and AAY90683-Y90687), and to DNA encoding them (AAA30709-A30743

CC and AAA30775-A30779). The mutant proteins of the invention contain a

CC mutation in a portion of the protein comprising intracellular loop 3

CC (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X,

CC is substituted for an endogenous residue in IC3 at a position 16 amino

CC acids N-terminal of an endogenous proline in TM6 to form a sequence

CC X-(AA)15-Pro. The endogenous amino acid is selected from Lys, His, Arg

CC or Ala, and is preferably Lys. When the endogenous residue at this

CC position is Lys, this residue is replaced by His, Arg or preferably Ala.

CC The 15 amino acid stretch between the substituted amino acid and the Pro

CC may be endogenous, non-endogenous, or a mixture of endogenous and

CC non-endogenous residues. The constitutively active GPCRs are useful for

CC identifying antagonists, agonists and partial agonists for use as

CC pharmaceutical agents. The mutant proteins are also useful in research

CC settings for elucidating the roles of the receptors in normal and

CC diseased conditions. Antagonists for a particular GPCR are useful for

CC treating diseases and disorders associated with that receptor. Because

CC the novel mutant GPCRs are constitutively active, they can be used

CC directly for screening of compounds without the need for endogenous

CC ligands. The present sequence represents cDNA encoding a human wild-type

CC GPCR used in an exemplification of the invention. This was cloned and

CC subjected to site-directed mutagenesis (SDM) to generate DNA encoding

CC the corresponding mutant of the invention.

XX Sequence 1164 BP; 246 A; 337 C; 288 G; 293 T; 0 other;

XX Query Match 3.1%; Score 32; DB 21; Length 1164;

XX Best Local Similarity 100.0%; Pred. No. 8.9e-06;

XX Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 790 CTCAGCTTCACCTACATGAACAGCATGCTGGA 821
|||||

Db 838 CTCAGCTTCACCTACATGAACAGCATGCTGGA 869

RESULT 13

AAA30738

ID AAA30738 standard; DNA; 1164 BP.

XX AAA30738;

AC AAA30738;

XX 21-AUG-2000 (first entry)

DT DNA encoding human mutant G protein-coupled receptor HM74 (I230K).

XX G protein-coupled receptor; GPCR; constitutively active;

XX intracellular loop 3; transmembrane domain 6; drug screening;

XX agonist; antagonist; mutant; ss.

XX Homo sapiens.

OS Synthetic.

OS WO200022129-A1.

PN 20-APR-2000.

XX 12-OCT-1999; 99WO-US23938.

XX 13-OCT-1998; 98US-0170496.

XX (AREN-) ARENA PHARM INC.

XX Behan DP, Chalmers DT, Liaw CW;

XX WPI: 2000-329165/28.

XX P-PSDB; AAY90672.

DR Non-endogenous constitutively activated human G protein-coupled

DR receptors, useful for identifying agonists for use as pharmaceutical

DR agents -

XX Example 2; Page 285-286; 341pp; English.

XX The invention relates to constitutively active, non-endogenous versions

XX of endogenous human orphan G protein-coupled receptors (GPCRs, AAY90643-

XX AAY90677 and AAY90683-Y90687), and to DNA encoding them (AAA30709-A30743

XX and AAA30775-A30779). The mutant proteins of the invention contain a

XX mutation in a portion of the protein comprising intracellular loop 3

XX (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X,

XX is substituted for an endogenous residue in IC3 at a position 16 amino

XX acids N-terminal of an endogenous proline in TM6 to form a sequence

XX X-(AA)15-Pro. The endogenous amino acid is selected from Lys, His, Arg

XX or Ala, and is preferably Lys. When the endogenous residue at this

XX position is Lys, this residue is replaced by His, Arg or preferably Ala.

XX The 15 amino acid stretch between the substituted amino acid and the Pro

XX may be endogenous, non-endogenous, or a mixture of endogenous and

XX non-endogenous residues. The constitutively active GPCRs are useful for

XX identifying antagonists, agonists and partial agonists for use as

XX pharmaceutical agents. The mutant proteins are also useful in research

XX settings for elucidating the roles of the receptors in normal and

XX diseased conditions. Antagonists for a particular GPCR are useful for

XX treating diseases and disorders associated with that receptor. Because

XX the novel mutant GPCRs are constitutively active, they can be used

XX directly for screening of compounds without the need for endogenous

XX ligands. The present sequence represents cDNA encoding a human wild-type

XX GPCR used in an exemplification of the invention. This was cloned and

XX subjected to site-directed mutagenesis (SDM) to generate DNA encoding

XX the corresponding mutant of the invention.

XX Sequence 1164 BP; 248 A; 335 C; 289 G; 292 T; 0 other;

XX Query Match 3.1%; Score 32; DB 21; Length 1164;

XX Best Local Similarity 100.0%; Pred. No. 8.9e-06;

XX Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 790 CTCAGCTTCACCTACATGACAGCAGCATGCTGGA 821
|||||
Db 838 CTCAGCTTCACCTACATGACAGCAGCATGCTGGA 869

RESULT 14

AAH51053
ID AAX16671 standard; cDNA; 1361 BP.

XX AC AAX16671;

XX DT 29-APR-1999 (first entry)

XX DE G-protein coupled receptor HM74A encoding cDNA.

XX KW HM74A receptor; G-protein coupled receptor; infection; pain; cancer;
KW diabetes; obesity; neurological disorder; heart failure; hypertension;
KW asthma; allergy; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
FT CDS 61..1152

FT FT /*tag= a
/product= "HM74A receptor"

XX PN W09856820-A1.

XX PD 17-DEC-1998.

XX PF 12-JUN-1998; 98WO-US12386.

XX PR 12-JUN-1997; 97US-0049480.

XX PA (SMIK) SMITHKLINE BEECHAM CORP.

XX PI Bergsma DJ, Elshourbagy NA, Guerrera SF, Li X, Mooney JL;

XX DR WPI; 1999-095273/08.

XX DR P-PSDB; AAW94654.

XX PT New isolated G-protein coupled receptor, HM74A - used to develop
PT products for treating e.g. infections, pain, cancers, diabetes,
PT obesity, neurological disorders, heart failure, hypertension, asthma
PT or allergies

XX PS Claim 2; Page 30-31; 40pp; English.

XX CC The present sequence encodes a member of the G-protein coupled receptor
CC (7TM receptor) family, designated the HM74A receptor. The proteins,
CC agonists, antagonists and polynucleotides can be used for treating
CC disorders associated with increased or reduced expression or activity
CC of HM74A, e.g. bacterial, fungal, protozoan and viral infections,
CC particularly infections caused by HIV-1 or HIV-2, pain, cancers,
CC diabetes, obesity, anorexia, bulimia, asthma, Parkinson's disease,
CC acute heart failure, hypotension, hypertension, urinary retention,
CC osteoporosis, angina pectoris, myocardial infarction, stroke, ulcers,
CC asthma, allergies, benign prostatic hypertrophy, migraine, vomiting,
CC psychotic and neurological disorders, including anxiety, schizophrenia,
CC manic depression, depression, delirium, dementia, and severe mental
CC retardation, and dyskinesias such as Huntington's disease or
CC Gilles de la Tourette's syndrome. The products can also be used for
CC detection, diagnosis and drug screening.

XX SQ Sequence 1361 BP; 291 A; 390 C; 342 G; 338 T; 0 other;

Query Match 3.1%; Score 32; DB 20; Length 1361;
Best Local Similarity 100.0%; Pred. No. 8.9e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 790 CTCAGCTTCACCTACATGACAGCAGCATGCTGGA 821

|||||

Db 838 CTCAGCTTCACCTACATGACAGCAGCATGCTGGA 929

RESULT 15

AAH51053

ID AAH51053 standard; DNA; 24 BP.

XX AC AAH51053;

XX DT 28-AUG-2001 (first entry)

XX DE Oligonucleotide SEQ ID 133.

XX KW G-protein-coupled receptor; nGPCR; seven transmembrane receptor;
KW signal transduction; schizophrenia; thyroid disorder; renal failure;
KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
KW cardiovascular disease; proliferative disorder; hormonal disorder;
KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
KW attention deficit-hyperactivity disorder/attention deficit disorder;
KW Parkinson's disease; migraine; senile dementia; inflammatory disease;
KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;
KW neuroprotective; ss.

XX OS Synthetic.

XX PN W0200136473-A2.

XX PD 25-MAY-2001.

XX PF 16-NOV-2000; 2000WO-US31581.

XX PR 16-NOV-1999; 99US-0165838.

XX PR 17-NOV-1999; 99US-0166071.

XX PR 19-NOV-1999; 99US-0166678.

XX PR 28-DEC-1999; 99US-0173396.

XX PR 22-FEB-2000; 2000US-0184129.

XX PR 28-FEB-2000; 2000US-0185421.

XX PR 02-MAR-2000; 2000US-0185534.

XX PR 03-MAR-2000; 2000US-0186811.

XX PR 09-MAR-2000; 2000US-0188114.

XX PR 17-MAR-2000; 2000US-0190310.

XX PR 21-MAR-2000; 2000US-0190800.

XX PR 20-APR-2000; 2000US-0198568.

XX PR 02-MAY-2000; 2000US-0201190.

XX PR 08-MAY-2000; 2000US-0203111.

XX PR 25-MAY-2000; 2000US-0207094.

XX PA (PHAA) PHARMACIA & UPJOHN CO.

XX PI Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;
PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;

XX DR WPI; 2001-389826/41.

XX PS Disclosure; Page 246; 261pp; English.

XX CC The present invention relates to novel G protein-coupled receptors
CC (nGPCR; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
CC 28, 31-38, 40, 41, 53-60) and their coding sequences (see
CC AAH50969-AAH51015 and AAH51105 and AAG80929-AAG80975 and AAG80977). The
CC present sequence is an oligonucleotide, which was used in the present
CC invention. GPCRs are also known as seven transmembrane receptors and
CC function in signal transduction. The nGPCR coding sequences are useful
CC for screening a human to diagnose a disorder affecting the brain or a
CC genetic predisposition, specifically schizophrenia. nGPCR are useful
CC for identifying compounds useful for treating schizophrenia. Detection of
CC nGPCR in a sample is useful as a diagnostic tool for diseases or
CC disorders e.g. thyroid disorders, renal failure, rheumatoid arthritis,
CC CNS disorders, infections such as HIV-1, metabolic and cardiovascular
CC diseases, proliferative disorders and hormonal disorders. Modulators of

CC nPCR activity have the utility for treating neurological disorders,
 CC including schizophrenia, ADHD/ADD (attention deficit-hyperactivity
 CC disorder/attention deficit disorder), and neuronal disorders such as
 CC Alzheimer's disease, Parkinson's disease, migraine and senile dementia.
 CC Additional disorders include inflammatory conditions (e.g. Crohn's
 CC disease), rheumatoid arthritis, autoimmune disorders, cancers,
 CC respiratory ailments such as asthma, and inflammatory diseases e.g.
 CC inflammatory bowel disease.
 XX
 SQ Sequence 24 BP; 3 A; 7 C; 5 G; 9 T; 0 other;
 Query Match 2.3%; Score 24; DB 22; Length 24;
 Best Local Similarity 100.0%; Pred. No. 0.094;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 707 GCGTGTCTGCTAGACTCTATTTC 730
 Db 1 GCGTGTCTGCTAGACTCTATTTC 24
 RESULT 16
 AAH51082
 ID AAH51082 standard; DNA; 24 BP.
 XX
 AC AAH51082;
 XX
 DT 28-AUG-2001 (first entry)
 XX
 DE Human nPCR11 PCR primer LW1660.
 XX
 KW G protein-coupled receptor; nPCR; seven transmembrane receptor;
 KW signal transduction; schizophrenia; thyroid disorder; renal failure;
 KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
 KW cardiovascular disease; proliferative disorder; hormonal disease;
 KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
 KW attention deficit-hyperactivity disorder/attention deficit disorder;
 KW Parkinson's disease; migraine; senile dementia; inflammatory disease;
 KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;
 KW neuroprotective; PCR primer; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200136473-A2.
 PN
 PD 25-MAY-2001.
 XX
 XX 16-NOV-2000; 2000WO-US31581.
 PF
 XX 16-NOV-1999; 99US-0165838.
 PR 17-NOV-1999; 99US-0166071.
 PR 19-NOV-1999; 99US-0166678.
 PR 28-DEC-1999; 99US-0173396.
 PR 22-FEB-2000; 2000US-0184129.
 PR 28-FEB-2000; 2000US-0185421.
 PR 28-FEB-2000; 2000US-0185554.
 PR 02-MAR-2000; 2000US-0186530.
 PR 03-MAR-2000; 2000US-0186811.
 PR 09-MAR-2000; 2000US-0188114.
 PR 17-MAR-2000; 2000US-0190310.
 PR 21-MAR-2000; 2000US-0190800.
 PR 20-APR-2000; 2000US-0198568.
 PR 02-MAY-2000; 2000US-0201190.
 PR 08-MAY-2000; 2000US-0203111.
 PR 25-MAY-2000; 2000US-0207094.
 XX
 XX (PHAA) PHARMACIA & UPJOHN CO.
 PA
 XX Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;
 PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;
 PI
 XX WPI; 2001-389826/41.
 DR
 XX
 XX New G protein-coupled receptor (nPCR-x) and its encoding

PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -
 XX
 PS Example 11; Page 143; 261pp; English.
 XX
 CC The present invention relates to novel G protein-coupled receptors
 CC (nPCR; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
 CC 28, 31-38, 40, 41, 53-60) and their coding sequences (see
 CC AAH50969-AAH51015 and AAH51105 and AAG80929-AAG80975 and AAG80977). The
 CC present sequence is a PCR primer, which was used in an example from the
 CC present invention. GPCRs are also known as seven transmembrane receptors
 CC and function in signal transduction. The nPCR coding sequences are
 CC useful for screening a human to diagnose a disorder affecting the brain
 CC or a genetic predisposition, specifically schizophrenia. nPCR are
 CC useful for identifying compounds useful for treating schizophrenia.
 CC Detection of nPCR in a sample is useful as a diagnostic tool for
 CC diseases or disorders e.g. thyroid disorders, renal failure, rheumatoid
 CC arthritis, CNS disorders, infections such as HIV-1, metabolic and
 CC cardiovascular diseases, proliferative disorders and hormonal disorders.
 CC Modulators of nPCR activity have the utility for treating neurological
 CC disorders, including schizophrenia, ADHD/ADD (attention deficit-
 CC hyperactivity disorder/attention deficit disorder), and neuronal
 CC disorders such as Alzheimer's disease, Parkinson's disease, migraine and
 CC senile dementia. Additional disorders include inflammatory conditions
 CC (e.g. Crohn's disease), rheumatoid arthritis, autoimmune disorders,
 CC cancers, respiratory ailments such as asthma, and inflammatory diseases
 CC e.g. inflammatory bowel disease.
 XX
 SQ Sequence 24 BP; 3 A; 7 C; 5 G; 9 T; 0 other;
 Query Match 2.3%; Score 24; DB 22; Length 24;
 Best Local Similarity 100.0%; Pred. No. 0.094;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 707 GCGTGTCTGCTAGACTCTATTTC 730
 Db 1 GCGTGTCTGCTAGACTCTATTTC 24
 RESULT 17
 AAH51050
 ID AAH51050 standard; DNA; 52 BP.
 XX
 AC AAH51050;
 XX
 DT 28-AUG-2001 (first entry)
 XX
 DE Oligonucleotide SEQ ID 130.
 XX
 KW G protein-coupled receptor; nPCR; seven transmembrane receptor;
 KW signal transduction; schizophrenia; thyroid disorder; renal failure;
 KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
 KW cardiovascular disease; proliferative disorder; hormonal disease;
 KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
 KW attention deficit-hyperactivity disorder/attention deficit disorder;
 KW Parkinson's disease; migraine; senile dementia; inflammatory disease;
 KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;
 KW neuroprotective; ss.
 XX
 OS Synthetic.
 XX
 PN WO200136473-A2.
 XX
 PD 25-MAY-2001.
 XX
 XX 16-NOV-2000; 2000WO-US31581.
 PF
 XX 16-NOV-1999; 99US-0165838.
 PR 17-NOV-1999; 99US-0166071.
 PR 19-NOV-1999; 99US-0166678.
 PR 28-DEC-1999; 99US-0173396.
 PR 22-FEB-2000; 2000US-0184129.
 PR 28-FEB-2000; 2000US-0185421.
 PR 28-FEB-2000; 2000US-0185554.
 PR 02-MAR-2000; 2000US-0186530.
 PR 03-MAR-2000; 2000US-0186811.
 PR 09-MAR-2000; 2000US-0188114.
 PR 17-MAR-2000; 2000US-0190310.
 PR 21-MAR-2000; 2000US-0190800.
 PR 20-APR-2000; 2000US-0198568.
 PR 02-MAY-2000; 2000US-0201190.
 PR 08-MAY-2000; 2000US-0203111.
 PR 25-MAY-2000; 2000US-0207094.
 XX
 XX (PHAA) PHARMACIA & UPJOHN CO.
 PA
 XX Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;
 PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;
 PI
 XX WPI; 2001-389826/41.
 DR
 XX
 XX New G protein-coupled receptor (nPCR-x) and its encoding

PR 02-MAR-2000; 2000US-0186530.
 PR 03-MAR-2000; 2000US-0186811.
 PR 09-MAR-2000; 2000US-0188114.
 PR 17-MAR-2000; 2000US-0190310.
 PR 21-MAR-2000; 2000US-0190800.
 PR 20-APR-2000; 2000US-0198568.
 PR 02-MAY-2000; 2000US-0201190.
 PR 08-MAY-2000; 2000US-0203111.
 PR 25-MAY-2000; 2000US-0207094.
 XX (PHAA) PHARMACIA & UPJOHN CO.
 XX Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;
 PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;
 XX WPI; 2001-389826/41.
 DR
 XX New G protein-coupled receptor (ngPCR-x) and its encoding
 PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -
 PT Disclosure; Page 245; 261pp; English.
 XX
 CC The present invention relates to novel G protein-coupled receptors
 CC (ngPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
 CC 28, 31-38, 40, 41, 53-60) and their coding sequences (see
 CC AAH50969-AAH51015 and AAH51105 and AAG80929-AAG80975 and AAG80977). The
 CC present sequence is an oligonucleotide, which was used in the present
 CC invention. GPCRs are also known as seven transmembrane receptors and
 CC function in signal transduction. The ngPCRx coding sequences are useful
 CC for screening a human to diagnose a disorder affecting the brain or a
 CC genetic predisposition, specifically schizophrenia. ngPCRx are useful
 CC for identifying compounds useful for treating schizophrenia. Detection of
 CC ngPCRx in a sample is useful as a diagnostic tool for diseases or
 CC disorders e.g. thyroid disorders, renal failure, rheumatoid arthritis,
 CC CNS disorders, infections such as HIV-1, metabolic and cardiovascular
 CC diseases, proliferative disorders and hormonal disorders. Modulators of
 CC ngPCRx activity have the utility for treating neurological disorders,
 CC including schizophrenia, ADHD/ADD (attention deficit-hyperactivity
 CC disorder/attention deficit disorder), and neuronal disorders such as
 CC Alzheimer's disease, Parkinson's disease, migraine and senile dementia.
 CC Additional disorders include inflammatory conditions (e.g. Crohn's
 CC disease), rheumatoid arthritis, autoimmune disorders, cancers,
 CC respiratory ailments such as asthma, and inflammatory diseases e.g.
 CC inflammatory bowel disease.
 XX
 SQ Sequence 52 BP; 12 A; 14 C; 12 G; 14 T; 0 other;
 Query Match 2.38; Score 24; DB 22; Length 52;
 Best Local Similarity 100.0%; Pred. No. 0.093;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 707 GCGTGTCTGCTAGACTCTATTTC 730
 Db 29 GCGTGTCTGCTAGACTCTATTTC 52
 RESULT 18
 AAH51079
 ID AAH51079 standard; DNA; 52 BP.
 XX
 AC AAH51079;
 XX
 DT 28-AUG-2001 (first entry)
 XX
 DE Human ngPCR11 PCR primer LW1658.
 XX
 KW G protein-coupled receptor; ngPCR; seven transmembrane receptor;
 KW signal transduction; schizophrenia; thyroid disorder; renal failure;
 KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
 KW cardiovascular disease; proliferative disorder; hormonal disorder;
 KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
 KW attention deficit-hyperactivity disorder; Alzheimer's disease; cancer;
 KW Parkinson's disease; migraine; senile dementia; inflammatory disease;

KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;
 KW neuroprotective; PCR primer; ss.
 OS Homo sapiens.
 XX WO200136473-A2.
 XX 25-MAY-2001.
 XX 16-NOV-2000; 2000WO-US31581.
 XX 16-NOV-1999; 99US-0165838.
 PR 17-NOV-1999; 99US-0166071.
 PR 19-NOV-1999; 99US-0166678.
 PR 28-DEC-1999; 99US-0173396.
 PR 22-FEB-2000; 2000US-0184129.
 PR 28-FEB-2000; 2000US-0185421.
 PR 28-FEB-2000; 2000US-0185554.
 PR 02-MAR-2000; 2000US-0186530.
 PR 03-MAR-2000; 2000US-0186811.
 PR 09-MAR-2000; 2000US-0188114.
 PR 17-MAR-2000; 2000US-0190310.
 PR 21-MAR-2000; 2000US-0190800.
 PR 20-APR-2000; 2000US-0198568.
 PR 02-MAY-2000; 2000US-0201190.
 PR 08-MAY-2000; 2000US-0203111.
 PR 25-MAY-2000; 2000US-0207094.
 XX (PHAA) PHARMACIA & UPJOHN CO.
 PA Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;
 PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;
 XX WPI; 2001-389826/41.
 DR
 XX New G protein-coupled receptor (ngPCR-x) and its encoding
 PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -
 PT Example 11; Page 143; 261pp; English.
 XX
 CC The present invention relates to novel G protein-coupled receptors
 CC (ngPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
 CC 28, 31-38, 40, 41, 53-60) and their coding sequences (see
 CC AAH50969-AAH51015 and AAH51105 and AAG80929-AAG80975 and AAG80977). The
 CC present sequence is a PCR primer, which was used in an example from the
 CC present invention. GPCRs are also known as seven transmembrane receptors
 CC and function in signal transduction. The ngPCRx coding sequences are
 CC useful for screening a human to diagnose a disorder affecting the brain
 CC or a genetic predisposition, specifically schizophrenia. ngPCRx are
 CC useful for identifying compounds useful for treating schizophrenia.
 CC Detection of ngPCRx in a sample is useful as a diagnostic tool for
 CC diseases or disorders e.g. thyroid disorders, renal failure, rheumatoid
 CC arthritis, CNS disorders, infections such as HIV-1, metabolic and
 CC cardiovascular diseases, proliferative disorders and hormonal disorders.
 CC Modulators of ngPCRx activity have the utility for treating neurological
 CC disorders, including schizophrenia, ADHD/ADD (attention deficit disorder),
 CC hyperactivity disorder/attention deficit disorder), and neuronal
 CC disorders such as Alzheimer's disease, Parkinson's disease, migraine and
 CC senile dementia. Additional disorders include inflammatory conditions
 CC (e.g. Crohn's disease), rheumatoid arthritis, autoimmune disorders,
 CC cancers, respiratory ailments such as asthma, and inflammatory diseases
 CC e.g. inflammatory bowel disease.
 XX
 SQ Sequence 52 BP; 12 A; 14 C; 12 G; 14 T; 0 other;
 Query Match 2.38; Score 24; DB 22; Length 52;
 Best Local Similarity 100.0%; Pred. No. 0.093;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 707 GCGTGTCTGCTAGACTCTATTTC 730
 Db 29 GCGTGTCTGCTAGACTCTATTTC 52

RESULT 19
AAH51029/c
ID AAH51029 standard; DNA; 33 BP.
XX
AC AAH51029;
XX
DT 28-AUG-2001 (first entry)
XX
DE Human nGPCR11 PCR primer LW1565.
XX
KW G protein-coupled receptor; nGPCR; seven transmembrane receptor;
KW signal transduction; schizophrenia; thyroid disorder; renal failure;
KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
KW cardiovascular disease; proliferative disorder; hormonal disorder;
KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
KW attention deficit-hyperactivity disorder; attention deficit disorder;
KW Parkinson's disease; migraine; senile dementia; inflammatory disease;
KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;
KW neuroprotective; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200136473-A2.
XX
PD 25-MAY-2001.
XX
PF 16-NOV-2000; 2000WO-US31581.
XX
PR 16-NOV-1999; 99US-0165838.
PR 17-NOV-1999; 99US-0166071.
PR 19-NOV-1999; 99US-0166678.
PR 28-DEC-1999; 99US-0173396.
PR 22-FEB-2000; 2000US-0184129.
PR 28-FEB-2000; 2000US-0185421.
PR 28-FEB-2000; 2000US-0185554.
PR 02-MAR-2000; 2000US-0186530.
PR 03-MAR-2000; 2000US-0186811.
PR 09-MAR-2000; 2000US-0188114.
PR 17-MAR-2000; 2000US-0190310.
PR 21-MAR-2000; 2000US-0190800.
PR 20-APR-2000; 2000US-0198568.
PR 02-MAY-2000; 2000US-0201190.
PR 08-MAY-2000; 2000US-0203111.
PR 25-MAY-2000; 2000US-0207094.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
PI Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;
PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;
XX
DR WPI; 2001-389826/41.
XX
XX New G protein-coupled receptor (nGPCR-x) and its encoding
PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -
XX
PS Example 2; Page 101; 261pp; English.
XX
CC The present invention relates to novel G protein-coupled receptors
CC (nGPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
CC 28, 31-38, 40, 41, 53-60) and their coding sequences (see
CC AAH50969-AAH51015 and AAH51105 and AAG80929-AAG80975 and AAG80977). The
CC present sequence is a PCR primer, which was used in an example from the
CC present invention. GPCRs are also known as seven transmembrane receptors
CC and function in signal transduction. The nGPCRx coding sequences are
CC useful for screening a human to diagnose a disorder affecting the brain
CC or a genetic predisposition, specifically schizophrenia. nGPCRx are
CC useful for identifying compounds useful for treating schizophrenia.
CC Detection of nGPCRx in a sample is useful as a diagnostic tool for
CC diseases or disorders e.g. thyroid disorders, renal failure, rheumatoid
CC arthritis, CNS disorders, infections such as HIV-1, metabolic and
CC cardiovascular diseases, proliferative disorders and hormonal disorders.
CC Modulators of nGPCRx activity have the utility for treating neurological

CC disorders, including schizophrenia, ADHD/ADD (attention deficit-
CC hyperactivity disorder/attention deficit disorder), and neuronal
CC disorders such as Alzheimer's disease, Parkinson's disease, migraine and
CC senile dementia. Additional disorders include inflammatory conditions
CC (e.g. Crohn's disease), rheumatoid arthritis, autoimmune disorders,
CC cancers, respiratory ailments such as asthma, and inflammatory diseases
CC e.g. inflammatory bowel disease.
XX
SQ Sequence 33 BP; 9 A; 8 C; 8 G; 8 T; 0 other;
Query Match 2.2%; Score 23; DB 22; Length 33;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1019 CCCACATTGTTGAGTGCGACTGA 1041
Db 33 CCCACATTGTTGAGTGCGACTGA 11
|||||
RESULT 20
AAH51028
ID AAH51028 standard; DNA; 33 BP.
XX
AC AAH51028;
XX
DT 28-AUG-2001 (first entry)
XX
DE Human nGPCR11 PCR primer LW1564.
XX
KW G protein-coupled receptor; nGPCR; seven transmembrane receptor;
KW signal transduction; schizophrenia; thyroid disorder; renal failure;
KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
KW cardiovascular disease; proliferative disorder; hormonal disorder;
KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
KW attention deficit-hyperactivity disorder; attention deficit disorder;
KW Parkinson's disease; migraine; senile dementia; inflammatory disease;
KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;
KW neuroprotective; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200136473-A2.
XX
PD 25-MAY-2001.
XX
PF 16-NOV-2000; 2000WO-US31581.
XX
PR 16-NOV-1999; 99US-0165838.
PR 17-NOV-1999; 99US-0166071.
PR 19-NOV-1999; 99US-0166678.
PR 28-DEC-1999; 99US-0173396.
PR 22-FEB-2000; 2000US-0184129.
PR 28-FEB-2000; 2000US-0185421.
PR 28-FEB-2000; 2000US-0185554.
PR 02-MAR-2000; 2000US-0186530.
PR 03-MAR-2000; 2000US-0186811.
PR 09-MAR-2000; 2000US-0188114.
PR 17-MAR-2000; 2000US-0190310.
PR 21-MAR-2000; 2000US-0190800.
PR 20-APR-2000; 2000US-0198568.
PR 02-MAY-2000; 2000US-0201190.
PR 08-MAY-2000; 2000US-0203111.
PR 25-MAY-2000; 2000US-0207094.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
PI Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;
PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;
XX
DR WPI; 2001-389826/41.
XX
XX New G protein-coupled receptor (nGPCR-x) and its encoding
PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -

```

XX Example 2; Page 101; 261pp; English.
PS
CC The present invention relates to novel G protein-coupled receptors
CC (nGPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
CC 28, 31-38, 40, 41, 53-60) and their coding sequences (see
CC AAH50969-AAH51015 and AAH51105 and AAG80929-AAG80975 and AAG80977). The
CC present sequence is a PCR primer, which was used in an example from the
CC present invention. GPCRs are also known as seven transmembrane receptors
CC and function in signal transduction. The nGPCRx coding sequences are
CC useful for screening a human to diagnose a disorder affecting the brain
CC or a genetic predisposition, specifically schizophrenia. nGPCRx are
CC useful for identifying compounds useful for treating schizophrenia.
CC Detection of nGPCRx in a sample is useful as a diagnostic tool for
CC diseases or disorders e.g. thyroid disorders, renal failure, rheumatoid
CC arthritis, CNS disorders, infections such as HIV-1, metabolic and
CC cardiovascular diseases, proliferative disorders and hormonal disorders.
CC Modulators of nGPCRx activity have the utility for treating neurological
CC disorders, including schizophrenia, ADHD/ADD (attention deficit-
CC hyperactivity disorder/attention deficit disorder), and neuronal
CC disorders such as Alzheimer's disease, Parkinson's disease, migraine and
CC senile dementia. Additional disorders include inflammatory conditions
CC (e.g. Crohn's disease), rheumatoid arthritis, autoimmune disorders,
CC cancers, respiratory ailments such as asthma, and inflammatory diseases
CC e.g. inflammatory bowel disease.
XX
SQ Sequence 33 BP; 7 A; 9 C; 9 G; 8 T; 0 other;

Query Match      2.0%; Score 21; DB 22; Length 33;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTACACGGGTCGTCTGC 21
   |||||
DB 13 ATGTACACGGGTCGTCTGC 33

RESULT 21
AAH51030
ID AAH51030 standard; DNA; 20 BP.
XX
AC AAH51030;
XX
DT 28-AUG-2001 (first entry)
XX
DE Human nGPCR11 PCR primer #1.
XX
KW G protein-coupled receptor; nGPCR; seven transmembrane receptor;
KW signal transduction; schizophrenia; thyroid disorder; renal failure;
KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
KW cardiovascular disease; proliferative disorder; hormonal disorder;
KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
KW attention deficit-hyperactivity disorder/attention deficit disorder;
KW Parkinson's disease; migraine; senile dementia; inflammatory disease;
KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;
KW neuroprotective; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200136473-A2.
XX
PD 25-MAY-2001.
XX
PF 16-NOV-2000; 2000WO-US31581.
XX
PR 16-NOV-1999; 99US-0165838.
PR 17-NOV-1999; 99US-0166071.
PR 19-NOV-1999; 99US-0166678.
PR 28-DEC-1999; 99US-0173396.
PR 22-FEB-2000; 2000US-0184129.
PR 28-FEB-2000; 2000US-0185421.
PR 28-FEB-2000; 2000US-0185554.
PR 02-MAR-2000; 2000US-0186530.

```

```

PR 03-MAR-2000; 2000US-0186811.
PR 09-MAR-2000; 2000US-0188114.
PR 17-MAR-2000; 2000US-0190310.
PR 21-MAR-2000; 2000US-0190800.
PR 20-APR-2000; 2000US-0198568.
PR 02-MAY-2000; 2000US-0201190.
PR 08-MAY-2000; 2000US-0203111.
PR 25-MAY-2000; 2000US-0207094.
XX (PHAA ) PHARMACIA & UPJOHN CO.
XX
XX Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;
PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;
XX WPI; 2001-399826/43.
XX
XX New G protein-coupled receptor (nGPCR-x) and its encoding
PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -
XX Example 4; Page 117; 261pp; English.
XX
CC The present invention relates to novel G protein-coupled receptors
CC (nGPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
CC 28, 31-38, 40, 41, 53-60) and their coding sequences (see
CC AAH50969-AAH51015 and AAH51105 and AAG80929-AAG80975 and AAG80977). The
CC present sequence is a PCR primer, which was used in an example from the
CC present invention. GPCRs are also known as seven transmembrane receptors
CC and function in signal transduction. The nGPCRx coding sequences are
CC useful for screening a human to diagnose a disorder affecting the brain
CC or a genetic predisposition, specifically schizophrenia. nGPCRx are
CC useful for identifying compounds useful for treating schizophrenia.
CC Detection of nGPCRx in a sample is useful as a diagnostic tool for
CC diseases or disorders e.g. thyroid disorders, renal failure, rheumatoid
CC arthritis, CNS disorders, infections such as HIV-1, metabolic and
CC cardiovascular diseases, proliferative disorders and hormonal disorders.
CC Modulators of nGPCRx activity have the utility for treating neurological
CC disorders, including schizophrenia, ADHD/ADD (attention deficit-
CC hyperactivity disorder/attention deficit disorder), and neuronal
CC disorders such as Alzheimer's disease, Parkinson's disease, migraine and
CC senile dementia. Additional disorders include inflammatory conditions
CC (e.g. Crohn's disease), rheumatoid arthritis, autoimmune disorders,
CC cancers, respiratory ailments such as asthma, and inflammatory diseases
CC e.g. inflammatory bowel disease.
XX
SQ Sequence 20 BP; 5 A; 7 C; 4 G; 4 T; 0 other;

Query Match      1.9%; Score 20; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 GAAGCCCGAGCACTGTTTACC 160
   |||||
DB 1 GAAGCCCGAGCACTGTTTACC 20

RESULT 22
AAH51031/c
ID AAH51031 standard; DNA; 20 BP.
XX
AC AAH51031;
XX
DT 28-AUG-2001 (first entry)
XX
DE Human nGPCR11 PCR primer #2.
XX
KW G protein-coupled receptor; nGPCR; seven transmembrane receptor;
KW signal transduction; schizophrenia; thyroid disorder; renal failure;
KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
KW cardiovascular disease; proliferative disorder; hormonal disorder;
KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
KW attention deficit-hyperactivity disorder/attention deficit disorder;
KW Parkinson's disease; migraine; senile dementia; inflammatory disease;
KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;

```

KW neuroprotective; PCR primer; ss.
 OS Homo sapiens.
 PN WO200136473-A2.
 XX 25-MAY-2001.
 PD XX
 PF 16-NOV-2000; 2000WO-US31581.
 XX
 XX 16-NOV-1999; 99US-0165838.
 PR 17-NOV-1999; 99US-0166071.
 PR 19-NOV-1999; 99US-0166678.
 PR 28-DEC-1999; 99US-0173396.
 PR 22-FEB-2000; 2000US-0184129.
 PR 28-FEB-2000; 2000US-0185421.
 PR 28-FEB-2000; 2000US-0185554.
 PR 02-MAR-2000; 2000US-0185530.
 PR 03-MAR-2000; 2000US-0186811.
 PR 09-MAR-2000; 2000US-0188114.
 PR 17-MAR-2000; 2000US-0190310.
 PR 21-MAR-2000; 2000US-0190800.
 PR 20-APR-2000; 2000US-0198568.
 PR 02-MAY-2000; 2000US-0201190.
 PR 08-MAY-2000; 2000US-0203111.
 PR 25-MAY-2000; 2000US-0207094.
 XX
 PA (PHAA) PHARMACIA & UPJOHN CO.
 XX
 XX Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;
 PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;
 XX WPI; 2001-389826/41.
 DR
 XX
 XX New G protein-coupled receptor (ngPCR-x) and its encoding
 PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -
 XX
 PS Example 4; Page 117; 261pp; English.
 XX
 CC The present invention relates to novel G protein-coupled receptors
 CC (ngPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
 CC 28, 31-38, 40, 41, 53-60) and their coding sequences (see
 CC AAH50969-AAH51015 and AAG80929-AAG80975 and AAG80977). The
 CC present sequence is a PCR primer, which was used in an example from the
 CC present invention. GPCRs are also known as seven transmembrane receptors
 CC and function in signal transduction. The ngPCRx coding sequences are
 CC useful for screening a human to diagnose a disorder affecting the brain
 CC or a genetic predisposition, specifically schizophrenia. ngPCRx are
 CC useful for identifying compounds useful for treating schizophrenia.
 CC Detection of ngPCRx in a sample is useful as a diagnostic tool for
 CC diseases or disorders e.g. thyroid disorders, renal failure, rheumatoid
 CC arthritis, CNS disorders, infections such as HIV-1, metabolic and
 CC cardiovascular diseases, proliferative disorders and hormonal disorders.
 CC Modulators of ngPCRx activity have the utility for treating neurological
 CC disorders, including schizophrenia, ADHD/ADD (attention deficit-
 CC hyperactivity disorder/attention deficit disorder), and neuronal
 CC disorders such as Alzheimer's disease, Parkinson's disease, migraine and
 CC senile dementia. Additional disorders include inflammatory conditions
 CC (e.g. Crohn's disease), rheumatoid arthritis, autoimmune disorders,
 CC cancers, respiratory ailments such as asthma, and inflammatory diseases
 CC e.g. inflammatory bowel disease.
 XX
 XX Sequence 20 BP; 5 A; 7 C; 4 G; 4 T; 0 other;
 SQ
 Query Match 1.9%; Score 20; DB 22; Length 20;
 Best Local Similarity 100.0%; Pred. No. 9.4;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 327 GGCTCGGACAGGTATTTC 346
 DB 20 GGCTCGGACAGGTATTTC 1

RESULT 23
 AAH51051/c
 ID AAH51051 standard; DNA; 20 BP.
 XX
 AC AAH51051;
 XX
 DT 28-AUG-2001 (first entry)
 XX
 DE Oligonucleotide SEQ ID 131.
 XX
 XX G protein-coupled receptor; ngPCR; seven transmembrane receptor;
 KW signal transduction; schizophrenia; thyroid disorder; renal failure;
 KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
 KW cardiovascular disease; proliferative disorder; hormonal disorder;
 KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
 KW attention deficit-hyperactivity disorder/attention deficit disorder;
 KW Parkinson's disease; migraine; senile dementia; inflammatory disease;
 KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;
 KW neuroprotective; ss.
 XX
 OS Synthetic.
 XX
 PN WO200136473-A2.
 XX
 PD 25-MAY-2001.
 XX
 XX 16-NOV-2000; 2000WO-US31581.
 XX
 XX 16-NOV-1999; 99US-0165838.
 PR 17-NOV-1999; 99US-0166071.
 PR 19-NOV-1999; 99US-0166678.
 PR 28-DEC-1999; 99US-0173396.
 PR 22-FEB-2000; 2000US-0184129.
 PR 28-FEB-2000; 2000US-0185421.
 PR 28-FEB-2000; 2000US-0185554.
 PR 02-MAR-2000; 2000US-0185530.
 PR 03-MAR-2000; 2000US-0186811.
 PR 09-MAR-2000; 2000US-0188114.
 PR 17-MAR-2000; 2000US-0190310.
 PR 21-MAR-2000; 2000US-0190800.
 PR 20-APR-2000; 2000US-0198568.
 PR 02-MAY-2000; 2000US-0201190.
 PR 08-MAY-2000; 2000US-0203111.
 PR 25-MAY-2000; 2000US-0207094.
 XX
 PA (PHAA) PHARMACIA & UPJOHN CO.
 XX
 XX Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;
 PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;
 XX WPI; 2001-389826/41.
 DR
 XX
 XX New G protein-coupled receptor (ngPCR-x) and its encoding
 PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -
 XX
 PS Disclosure; Page 245; 261pp; English.
 XX
 CC The present invention relates to novel G protein-coupled receptors
 CC (ngPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
 CC 28, 31-38, 40, 41, 53-60) and their coding sequences (see
 CC AAH50969-AAH51015 and AAG80929-AAG80975 and AAG80977). The
 CC present sequence is an oligonucleotide, which was used in the present
 CC invention. GPCRs are also known as seven transmembrane receptors and
 CC function in signal transduction. The ngPCRx coding sequences are useful
 CC for screening a human to diagnose a disorder affecting the brain or a
 CC genetic predisposition, specifically schizophrenia. ngPCRx are useful
 CC for identifying compounds useful for treating schizophrenia. Detection of
 CC ngPCRx in a sample is useful as a diagnostic tool for diseases or
 CC disorders e.g. thyroid disorders, renal failure, rheumatoid arthritis,
 CC CNS disorders, infections such as HIV-1, metabolic and cardiovascular
 CC diseases, proliferative disorders and hormonal disorders. Modulators of
 CC ngPCRx activity have the utility for treating neurological disorders,
 CC including schizophrenia, ADHD/ADD (attention deficit-hyperactivity

CC disorder/attention deficit disorder), and neuronal disorders such as
 CC Alzheimer's disease, Parkinson's disease, migraine and senile dementia.
 CC Additional disorders include inflammatory conditions (e.g. Crohn's
 CC disease), rheumatoid arthritis, autoimmune disorders, cancers,
 CC respiratory ailments such as asthma, and inflammatory diseases e.g.
 CC inflammatory bowel disease.

XX
 SQ Sequence 20 BP; 5 A; 8 C; 3 G; 4 T; 0 other;

Query Match 1.9%; Score 20; DB 22; Length 20;
 Best Local Similarity 100.0%; Pred. No. 9.4; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 962 GGAGTTCATCAGTGTGGCA 981
 |||||
 DB 20 GGAGTTCATCAGTGTGGCA 1

RESULT 24
 AAH51080/c
 ID AAH51080 standard; DNA: 20 BP.
 XX
 AC AAH51080;
 XX
 DT 28-AUG-2001 (first entry)
 DE Human nGPCR11 PCR primer LW1661.

XX G protein-coupled receptor; nGPCR; seven transmembrane receptor;
 KW signal transduction; schizophrenia; thyroid disorder; renal failure;
 KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
 KW cardiovascular disease; proliferative disorder; hormonal disorder;
 KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
 KW attention deficit-hyperactivity disorder/attention deficit disorder;
 KW Parkinson's disease; migraine; senile dementia; inflammatory disease;
 KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;
 KW neuroprotective; PCR primer; ss.
 XX Homo sapiens.
 OS
 XX
 PN WO200136473-A2.
 XX
 XX 25-MAY-2001.
 PD
 XX
 XX 16-NOV-2000; 2000WO-US31581.
 XX
 PR 16-NOV-1999; 99US-0165838.
 PR 17-NOV-1999; 99US-0166071.
 PR 19-NOV-1999; 99US-0166678.
 PR 28-DEC-1999; 99US-0173396.
 PR 22-FEB-2000; 2000US-0184129.
 PR 28-FEB-2000; 2000US-0185421.
 PR 28-FEB-2000; 2000US-0185554.
 PR 02-MAR-2000; 2000US-0186530.
 PR 03-MAR-2000; 2000US-0186811.
 PR 09-MAR-2000; 2000US-0188114.
 PR 17-MAR-2000; 2000US-0190310.
 PR 21-MAR-2000; 2000US-0190800.
 PR 20-APR-2000; 2000US-0198568.
 PR 02-MAY-2000; 2000US-0201190.
 PR 08-MAY-2000; 2000US-0203111.
 PR 25-MAY-2000; 2000US-0207094.
 XX
 XX (PHAA) PHARMACIA & UPJOHN CO.
 XX
 PI Vogell G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;
 PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;
 XX
 DR WPI; 2001-389826/41.
 XX
 XX New G protein-coupled receptor (nGPCR-x) and its encoding
 PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -
 XX

PS Example 11; Page 143; 261pp; English.
 XX
 CC The present invention relates to novel G protein-coupled receptors
 CC (nGPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
 CC 28, 31-38, 40, 41, 53-60) and their coding sequences (see
 CC AAH50969-AAH51015 and AAH51105 and AAG80929-AAG80975 and AAG80977). The
 CC present sequence is a PCR primer, which was used in an example from the
 CC present invention. nGPCRs are also known as seven transmembrane receptors
 CC and function in signal transduction. The nGPCRx coding sequences are
 CC useful for screening a human to diagnose a disorder affecting the brain
 CC or a genetic predisposition, specifically schizophrenia. nGPCRx are
 CC useful for identifying compounds useful for treating schizophrenia.
 CC Detection of nGPCRx in a sample is useful as a diagnostic tool for
 CC diseases or disorders e.g. thyroid disorders, renal failure, rheumatoid
 CC arthritis, CNS disorders, infections such as HIV-1, metabolic and
 CC cardiovascular diseases, proliferative disorders and hormonal disorders.
 CC Modulators of nGPCRx activity have the utility for treating neurological
 CC disorders, including schizophrenia, ADHD/ADD (attention deficit-
 CC hyperactivity disorder/attention deficit disorder), and neuronal
 CC disorders such as Alzheimer's disease, Parkinson's disease, migraine and
 CC senile dementia. Additional disorders include inflammatory conditions
 CC (e.g. Crohn's disease), rheumatoid arthritis, autoimmune disorders,
 CC cancers, respiratory ailments such as asthma, and inflammatory diseases
 CC e.g. inflammatory bowel disease.

XX
 SQ Sequence 20 BP; 5 A; 8 C; 3 G; 4 T; 0 other;

Query Match 1.9%; Score 20; DB 22; Length 20;
 Best Local Similarity 100.0%; Pred. No. 9.4; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 962 GGAGTTCATCAGTGTGGCA 981
 |||||
 DB 20 GGAGTTCATCAGTGTGGCA 1

RESULT 25
 AAH51052/c
 ID AAH51052 standard; DNA: 48' BP.
 XX
 AC AAH51052;
 XX
 DT 28-AUG-2001 (first entry)
 DE Oligonucleotide SEQ ID 132.
 XX
 XX G protein-coupled receptor; nGPCR; seven transmembrane receptor;
 KW signal transduction; schizophrenia; thyroid disorder; renal failure;
 KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
 KW cardiovascular disease; proliferative disorder; hormonal disorder;
 KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
 KW attention deficit-hyperactivity disorder/attention deficit disorder;
 KW Parkinson's disease; migraine; senile dementia; inflammatory disease;
 KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;
 KW neuroprotective; ss.
 XX
 OS Synthetic.
 XX
 XX WO200136473-A2.
 PN
 XX
 XX 25-MAY-2001.
 PD
 XX
 XX 16-NOV-2000; 2000WO-US31581.
 XX
 PR 16-NOV-1999; 99US-0165838.
 PR 17-NOV-1999; 99US-0166071.
 PR 19-NOV-1999; 99US-0166678.
 PR 28-DEC-1999; 99US-0173396.
 PR 22-FEB-2000; 2000US-0184129.
 PR 28-FEB-2000; 2000US-0185421.
 PR 28-FEB-2000; 2000US-0185554.
 PR 02-MAR-2000; 2000US-0186530.
 PR 03-MAR-2000; 2000US-0186811.

AAL37148/c
ID AAL37148 standard; DNA; 4715 BP.
AC AAL37148;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide SEQ ID NO 3513.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.
XX
OS Homo sapiens.
XX
PN WO200155367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01338.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225366.
PR 14-AUG-2000; 2000US-0225367.
PR 14-AUG-2000; 2000US-0225468.
PR 14-AUG-2000; 2000US-0225469.
PR 14-AUG-2000; 2000US-0225477.
PR 14-AUG-2000; 2000US-0225478.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249299.

PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 PI
 XX
 DR WPI; 2001-451937/48.
 XX
 XX Isolated polypeptide for treating, preventing and/ or prognosing
 PT disorders related to the musculoskeletal system including
 PT musculoskeletal cancers and also for testing and detection e.g.
 PT diagnosis -
 XX
 XX Example 2; SEQ ID NO 3513; 781pp + Sequence Listing; English.
 PS
 XX The invention relates to novel genes (AAL34669-AAL37666) and proteins
 CC (ABB03087-ABB04109) associated with the musculoskeletal system useful
 CC for preventing, treating or ameliorating medical conditions e.g. by
 CC protein or gene therapy. The genes are isolated from a range of human
 CC tissues disclosed in the specification. The nucleic acids, proteins,
 CC antibodies and (ant)agonists are useful in the diagnosis, treatment
 CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
 CC other cancers of the adrenal gland, bone, bone marrow, breast,
 CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
 CC (c) cardiovascular disorders such as myocardial ischaemias; (d) wound
 CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
 CC and (f) infectious diseases such as viral, bacterial, fungal and
 CC parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 4715 BP; 1286 A; 1092 C; 1086 G; 1251 T; 0 other;
 Query Match 1.9%; Score 20; DB 22; Length 4715;
 Best Local Similarity 100.0%; Pred. No. 9;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 189 CCTTATGATCTGCCTGCTT 208
 Db 3122 CCTTATGATCTGCCTGCTT 3103
 RESULT 28
 ABL17061/c
 ID ABL17061 standard; DNA; 912 BP.
 XX
 XX ABL17061;
 AC
 XX
 DT 26-MAR-2002 (first entry)
 XX
 XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 2656.
 DE
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ds.
 KW
 XX Drosophila melanogaster.
 OS

XX WO200171042-A2.
 PN
 XX
 PD 27-SEP-2001.
 XX
 XX 23-MAR-2001; 2001WO-US09231.
 PF
 XX 23-MAR-2000; 2000US-191637P.
 PR
 PR 11-JUL-2000; 2000US-0614150.
 PR
 XX (PEKE) PE CORP NY.
 PA
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI
 XX WPI; 2001-656860/75.
 DR
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PT
 XX
 PS Claim 1; SEQ ID NO 2656; 21pp + Sequence Listing; English.
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 912 BP; 265 A; 188 C; 225 G; 234 T; 0 other;
 Query Match 1.8%; Score 19; DB 23; Length 912;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 805 ATGAACAGCATGCTGGATC 823
 Db 41 ATGAACAGCATGCTGGATC 23
 RESULT 29
 ABL02109/c
 ID ABL02109 standard; cDNA; 2947 BP.
 XX
 XX ABL02109;
 AC
 XX
 DT 26-MAR-2002 (first entry)
 XX
 XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 809.
 DE
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 KW
 XX Drosophila melanogaster.
 OS
 XX WO200171042-A2.
 PN
 XX
 PD 27-SEP-2001.
 XX
 XX 23-MAR-2001; 2001WO-US09231.
 PF
 XX 23-MAR-2000; 2000US-191637P.
 PR
 PR 11-JUL-2000; 2000US-0614150.
 PR
 XX (PEKE) PE CORP NY.
 PA
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI
 XX

DR WPI; 2001-656860/75.
DR P-PSDB; ABB58006.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 809; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2947 BP; 772 A; 921 C; 719 G; 535 T; 0 other;

Query Match 1.8%; Score 19; DB 23; Length 2947;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 609 GAGCGGAGGAGCAGCGTG 627
Db 1942 GAGCGGAGGAGCAGCGTG 1924
|||||

RESULT 30
ABL17060/C
ID ABL17060 standard; DNA; 2977 BP.
XX
AC ABL17060;
XX
XX 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 2653.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
PS Claim 1; SEQ ID NO 2653; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2977 BP; 845 A; 616 C; 680 G; 836 T; 0 other;

Query Match 1.8%; Score 19; DB 23; Length 2977;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 805 ATGAACAGCATGCTGGATC 823
Db 1041 ATGAACAGCATGCTGGATC 1023
|||||

RESULT 31
ABL19418/C
ID ABL19418 standard; DNA; 3482 BP.
XX
AC ABL19418;
XX
XX 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 9727.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
PS Claim 1; SEQ ID NO 9727; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 3482 BP; 836 A; 797 C; 959 G; 890 T; 0 other;

Query Match 1.8%; Score 19; DB 23; Length 3482;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 805 ATGAACAGCATGCTGGATC 823


```
|||||
Db 3212 ATGAACAGCATGCTGGATC 3194

RESULT 32
AAS27822
ID AAS27822 standard; DNA; 10098 BP.
XX
AC AAS27822;
XX
DT 07-NOV-2001 (first entry)
XX
DE DNA encoding novel signal transduction pathway protein, Seq ID 1482.
XX
KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS; ds;
KW acquired immune deficiency syndrome.
XX
OS Homo sapiens.
XX
PN WO200154733-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01312.
XX
31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 12-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
```


Query Match 1.8%; Score 19; DB 22; Length 12619;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 851 CCTTCCCAAAATCTACAA 869
DB 1940 CCTTCCCAAAATCTACAA 1958
|||||

RESULT 34

ABL02108
ID ABL02108 standard; cDNA; 19175 BP.

XX AC ABL02108;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 806.

XX KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-191637P.

XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR P-PSDB; ABB58005.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

XX PS Claim 1; SEQ ID NO 806; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB5737-ABB72072).

XX CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 19175 BP; 5148 A; 4250 C; 4561 G; 5216 T; 0 other;

Query Match 1.8%; Score 19; DB 23; Length 19175;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 609 GAGCGGAGGCAGCAGCTG 627
|||||

DB 2006 GAGCGGAGGCAGCAGCTG 2024
|||||

RESULT 35

AAV09268/c

ID AAV09268 standard; cDNA; 433 BP.

XX

AC AAV09268;

XX DT 07-JUL-1998 (first entry)

XX DE Nucleotide sequence of the 5' portion of the C195_1 protein.

XX CC C195_1 protein; human PBMC cDNA library; cytokine activity; ss;
KW peripheral blood mononuclear cell; nutritional activity; homology;
KW cell proliferation/differentiation activity; EST; expressed sequence tag.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
CDS 13..433

XX FT /*tag= a

XX FT /product= "C195_1 protein"

XX FT /note= "no stop codon found at the 3' end"

XX PN WO9748801-A2.

XX PD 24-DEC-1997.

XX PF 16-JUN-1997; 97WO-US10501.

XX PR 17-JUN-1996; 96US-0664596.

XX XX (GEMY) GENETICS INST INC.

XX XX Bowman M, Evans C, Jacobs K, LaVallie ER, McCoy JM;

XX PI Merberg D, Racie LA, Spaulding V, Treacy M;

XX XX WPI; 1998-063142/06.

XX DR P-PSDB; AAW42074.

XX CC Poly:nucleotide(s) and proteins obtained from human PBMC, dendritic
PT cell, adult brain, foetal brain and adult testes cDNA libraries -
PT used in research, detection and therapy of, e.g. cytokine and cell
PT proliferation or differentiation

XX PS Claim 1; Page 47; 78pp; English.

XX CC This nucleotide sequence encodes the 5' portion of the C195_1 protein
CC which was isolated from a human PBMC cDNA library. The products of
CC the polynucleotides of the invention can be used in research, detection
CC and therapy, as they may have nutritional activity, cytokine and cell
CC proliferation/differentiation activity. A search against the Genbank
CC database demonstrated that this sequence has at least some homology
CC with two ESTs identified as "yg1906.r1 Homo sapiens cDNA clone
CC 31937 5'" and "ym15f12.r1 Homo sapiens cDNA clone 48025 5'".

XX SQ Sequence 433 BP; 125 A; 85 C; 98 G; 125 T; 0 other;

Query Match 1.7%; Score 18; DB 19; Length 433;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 531 TGACATCATGTTCCAGCT 548
|||||

DB 341 TGACATCATGTTCCAGCT 324
|||||

RESULT 36

AAV04273/c

ID AAV04273 standard; cDNA; 433 BP.

XX AC AAV04273;

XX DT 22-JUN-1998 (first entry)

XX DE Secreted protein C195_1 cDNA.

XX KW Secreted protein; C195_1; peripheral blood mononuclear cell; PBMC;
KW protein factor; human; ds.

```
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT mat_peptide 13..433
XX FT /*tag= a
XX PN WO9748800-A1.
XX PD 24-DEC-1997.
XX PF 16-JUN-1997; 97WO-US10500.
XX PR 25-OCT-1996; 96US-0738367.
XX PR 17-JUN-1996; 96US-0664596.
XX PR 27-SEP-1996; 96US-0721976.
XX PA (GEMY ) GENETICS INST INC.
XX PI Jacobs K, Lavallee ER, McCoy JM, Merberg D, Racie LA;
XX PI Spaulding V, Treacy M;
XX DR WPI; 1998-063141/06.
XX DR P-PSDB; AAW41600.
XX PT Nucleic acid encoding secreted protein from human peripheral blood
XX PT mononuclear cells - useful, e.g. as immunomodulators, antitumour
XX PT agents, promoters of tissue growth, haemostatic and thrombolytic
XX PT agents etc.
XX PS Claim 1; Page 35; 52pp; English.
XX CC This cDNA clone, designated C195_1, codes for a novel secreted
XX CC protein (see AAW41600) of human peripheral blood mononuclear cells
XX CC (PBMC). It was isolated from a PBMC cDNA library using methods
XX CC which are selective for cDNAs encoding secreted proteins. The 433
XX CC bp sequence represents the 5' part of the clone; the 3' portion of
XX CC C195_1, including the polyA tail, is provided in AAV04275. C195_1 is
XX CC deposited as ATCC 98079, a composite clone. An additional isolate
XX CC of C195_1, designated C195_4 (see AAV04274), is deposited as ATCC
XX CC 98192. Both sequences can be recovered using a probe (see AAV04276).
XX CC The isolated nucleic acid may be used to express recombinant secreted
XX CC proteins; as a tissue/molecular weight marker; for chromosome
XX CC identification; to identify possible genetic disorders; to isolate
XX CC new related DNA; as a source of PCR primers; to generate anti-protein
XX CC or anti-DNA antibodies; in interaction trap assays to identify
XX CC sequences that encode interacting proteins etc. The secreted
XX CC proteins can be used to screen compounds for biological activity;
XX CC to raise antibodies; as tissue markers; for isolation of related
XX CC receptors and ligands and as nutritional sources. They may also
XX CC have biological activities, e.g. cytokine, cell proliferation or
XX CC differentiation activity, immunosuppressant, immunostimulant,
XX CC regulation of haematopoiesis, modulation of fertility, chemotactic,
XX CC chemokinetic, haemostatic, thrombolytic, antimicrobials, analgesic,
XX CC antipsoriatic, etc. No evidence is given to support any of these
XX CC activities or applications.
XX SQ Sequence 433 BP; 125 A; 85 C; 98 G; 125 T; 0 other;

Query Match 1.7%; Score 18; DB 19; Length 433;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 531 TGACATCATGTTCCAGCT 548
DB 341 TGACATCATGTTCCAGCT 324

RESULT 37
AAI92474
ID AAI92474 standard; cDNA; 444 BP.
XX AC AAI92474;
```

```
XX 06-NOV-2001 (first entry)
XX Human polynucleotide SEQ ID NO 12534.
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation; ss.
XX Homo sapiens.
XX WO200164835-A2.
XX 07-SEP-2001.
XX 26-FEB-2001; 2001WO-US04927.
XX 28-FEB-2000; 2000US-0515126.
XX 18-MAY-2000; 2000US-0577409.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-514838/56.
XX P-PSDB; AAO12543.
XX Isolated nucleic acids and polypeptides, useful for preventing
XX diagnosing and treating e.g. leukaemia, inflammation and immune
XX disorders -
XX Claim 1; SEQ ID NO 12534; 1399pp + Sequence Listing; English.
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 444 BP; 88 A; 110 C; 108 G; 138 T; 0 other;

Query Match 1.7%; Score 18; DB 22; Length 444;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 422 CCCTGGTCATCCTCGGAA 439
DB 193 CCCTGGTCATCCTCGGAA 210

RESULT 38
AAV86048/c
ID AAV86048 standard; cDNA; 451 BP.
XX AC AAV86048;
XX 27-APR-1999 (first entry)
XX EST clone C195.
XX Expressed sequence tag; secreted protein; haematopoiesis regulator;
XX tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
XX chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
```

KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.

XX Homo sapiens.

XX WO9845435-A2.

XX 15-OCT-1998.

XX 10-APR-1998; 98WO-US06954.

XX 10-APR-1997; 97US-0835913.

XX (GEMY) GENETICS INST INC.

XX Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

XX Racie LA, Spaulding V, Treacy M;

XX WPI; 1999-070076/06.

XX New polynucleotides encoding human secreted proteins - derived from
PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
PT ovary, pituitary, retina and colon cDNA libraries

XX Claim 1; Page 101; 633pp; English.

XX This sequence represents an expressed sequence tag (EST), and is a
CC polynucleotide of the invention. The polynucleotides of the invention are
CC all secreted EST sequences isolated from a variety of human tissue
CC sources. The EST sequences and proteins encoded by them are predicted to
CC have useful biological activities which would make them suitable for
CC treating, preventing or ameliorating medical conditions in humans and
CC animals, although no supporting data is given. Suggested activities
CC include nutritional activity, immune stimulating or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
CC activity. The EST sequences are also stated to be useful for gene
CC therapy.

XX Sequence 451 BP; 128 A; 91 C; 102 G; 130 T; 0 other;

Query Match 1.7%; Score 18; DB 20; Length 451;

Best Local Similarity 100.0%; Pred. No. 92;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 531 TGACATCATGTTCCAGCT 548

DB 341 TGACATCATGTTCCAGCT 324

RESULT 39

AAI82739

ID AAI82739 standard; cDNA; 464 BP.

XX AAI82739;

XX 06-NOV-2001 (first entry)

XX Human polynucleotide SEQ ID NO 2799.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.

XX Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US04927.

XX

PR 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

DR P-PSDB; AAO02808.

XX Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -

XX Claim 1; SEQ ID NO 2799; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 464 BP; 124 A; 105 C; 118 G; 117 T; 0 other;

Query Match 1.7%; Score 18; DB 22; Length 464;

Best Local Similarity 100.0%; Pred. No. 92;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 774 GGCCCTGCACATAACCCCT 791

DB 171 GGCCCTGCACATAACCCCT 188

RESULT 40

AAS36762

ID AAS36762 standard; DNA; 477 BP.

XX AAS36762;

XX 17-DEC-2001 (first entry)

XX Human cardiovascular system antigen genomic DNA SEQ ID NO 2262.

XX Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;
KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
KW cerebroprotective; nontropic; antibacterial; virucide; fungicide; cancer;
KW ophthalmological; vulnerary; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; tissue regeneration;
KW anti-infertility.

XX Homo sapiens.

XX WO200155321-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01340.

PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225271.
PR 14-AUG-2000; 2000US-0225275.
PR 14-AUG-2000; 2000US-0225275.
PR 14-AUG-2000; 2000US-0225275.
PR 18-AUG-2000; 2000US-0225279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 03-SEP-2000; 2000US-0229509.
PR 03-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-024617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 03-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2000US-0254097.
PR 05-JAN-2001; 2000US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-451930/48.
DR
XX
PT New cardiovascular system related polynucleotides and polypeptides,
PT useful for diagnosing, treating and/or preventing disorders of the
PT cardiovascular system -

20-OCT-2000; 2000US-02411221.
20-OCT-2000; 2000US-02411785.
20-OCT-2000; 2000US-02411786.
20-OCT-2000; 2000US-02411787.
20-OCT-2000; 2000US-0241808.
20-OCT-2000; 2000US-0241809.
20-OCT-2000; 2000US-0241826.
01-NOV-2000; 2000US-0244617.
08-NOV-2000; 2000US-0246474.
08-NOV-2000; 2000US-0246475.
08-NOV-2000; 2000US-0246476.
08-NOV-2000; 2000US-0246477.
08-NOV-2000; 2000US-0246478.
08-NOV-2000; 2000US-0246523.
08-NOV-2000; 2000US-0246524.
08-NOV-2000; 2000US-0246525.
08-NOV-2000; 2000US-0246526.
08-NOV-2000; 2000US-0246527.
08-NOV-2000; 2000US-0246528.
08-NOV-2000; 2000US-0246532.
08-NOV-2000; 2000US-0246609.
08-NOV-2000; 2000US-0246610.
08-NOV-2000; 2000US-0246611.
08-NOV-2000; 2000US-0246613.
17-NOV-2000; 2000US-0249207.
17-NOV-2000; 2000US-0249208.
17-NOV-2000; 2000US-0249209.
17-NOV-2000; 2000US-0249210.
17-NOV-2000; 2000US-0249211.
17-NOV-2000; 2000US-0249212.
17-NOV-2000; 2000US-0249213.
17-NOV-2000; 2000US-0249214.
17-NOV-2000; 2000US-0249215.
17-NOV-2000; 2000US-0249216.
17-NOV-2000; 2000US-0249217.
17-NOV-2000; 2000US-0249218.
17-NOV-2000; 2000US-0249244.
17-NOV-2000; 2000US-0249245.
17-NOV-2000; 2000US-0249264.
17-NOV-2000; 2000US-0249265.
17-NOV-2000; 2000US-0249297.
17-NOV-2000; 2000US-0249299.
01-DEC-2000; 2000US-0250300.
01-DEC-2000; 2000US-0250310.
01-DEC-2000; 2000US-0250391.
05-DEC-2000; 2000US-0251030.
05-DEC-2000; 2000US-0251031.
05-DEC-2000; 2000US-0251988.
05-DEC-2000; 2000US-0256719.
06-DEC-2000; 2000US-0251479.
08-DEC-2000; 2000US-0251856.
08-DEC-2000; 2000US-0251868.
08-DEC-2000; 2000US-0251869.
08-DEC-2000; 2000US-0251989.
08-DEC-2000; 2000US-0251990.
11-DEC-2000; 2000US-0254097.
05-JAN-2001; 2001US-0259678.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-451930/48.
New cardiovascular system related polynucleotides and polypeptides,
useful for diagnosing, treating and/or preventing disorders of the
cardiovascular system -
Claim 1: SEQ ID No 2263; 674pp; English.
Sequences AAS35741-AAS36942 represent genomic DNA molecules, which encode
the cardiovascular system antigen polypeptides of the invention.
Cardiovascular system antigens and their associated polynucleotides are
useful in the diagnosis, treatment and prevention of various types of

disorders in e.g. humans, mice, rabbits, goats, cats, dogs, chickens or sheep. A pathological condition can be determined by detecting the presence or absence of a mutation in a cardiovascular system antigen polynucleotide. The treatable disorders include autoimmune diseases such as rheumatoid arthritis, hyperproliferative disorders such as neoplasms of the breast or liver, cardiovascular disorders such as cardiac arrest, cerebrovascular disorders such as cerebral ischaemia, nervous system disorders such as Alzheimer's disease, infections caused by bacteria, viruses and fungi, ocular disorders such as corneal infection, endocrine disorders such as premature labour and infertility, gastrointestinal disorders such as Crohn's disease, renal disorders such as glomerulonephritis and respiratory disorders such as asthma and pleurisy. The polypeptides can also be used to aid wound healing, to prevent skin aging due to sunburn, to maintain organs before transplantation, to regenerate tissues and in chemotaxis.
Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Query Match 1.7%; Score 18; DB 22; Length 477;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 894 ACCAAGCAGCCAGGACA 911
|||||
Db 265 ACCAAGCAGCCAGGACA 282

RESULT 42

AAS35641
ID AAS35641 standard; cDNA; 489 BP.

AC AAS35641;

XX 17-DEC-2001 (first entry)

XX Human cardiovascular system antigen cDNA polynucleotide SEQ ID No 526.

Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;
chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;
anti-rheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
cerebroprotective; neurotropic; antibacterial; virucide; fungicide; cancer;
ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;
hyperproliferative disorder; breast; liver; cardiovascular disorder; ss;
cerebrovascular disorder; nervous system disorder; bacterial infection;
fungal infection; viral infection; ocular disorder; endocrine disorder;
gastrointestinal disorder; renal disorder; respiratory disorder;
wound healing; skin aging; organ transplantation; tissue regeneration;
anti-infertility.

OS Homo sapiens.

XX WO200155321-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01340.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-02117496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 18-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0225799.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0234999.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236602.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0248477.
PR 08-NOV-2000; 2000US-0248478.
PR 08-NOV-2000; 2000US-0248524.
PR 08-NOV-2000; 2000US-0248525.
PR 08-NOV-2000; 2000US-0248526.
PR 08-NOV-2000; 2000US-0248527.
PR 08-NOV-2000; 2000US-0248528.
PR 08-NOV-2000; 2000US-0248532.
PR 08-NOV-2000; 2000US-0248609.
PR 08-NOV-2000; 2000US-0248610.
PR 08-NOV-2000; 2000US-0248611.
PR 08-NOV-2000; 2000US-0248613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-451930/48.
XX P-PSDB; AAU22367.
PT New cardiovascular system related polynucleotides and polypeptides,
PT useful for diagnosing, treating and/or preventing disorders of the
PT cardiovascular system -
XX
XX Claim 1; SEQ ID No 526; 674pp; English.
XX
CC Sequences AAS35126-AAS35740 represent cDNA molecules, which encode the
CC cardiovascular system antigen polypeptides of the invention.
CC Cardiovascular system antigens and their associated polynucleotides are
CC useful in the diagnosis, treatment and prevention of various types of
CC disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
CC chickens or sheep. A pathological condition can be determined by
CC detecting the presence or absence of a mutation in a cardiovascular
CC system antigen polynucleotide. The treatable disorders include autoimmune
CC diseases such as rheumatoid arthritis, hyperproliferative disorders such
CC as neoplasms of the breast or liver, cardiovascular disorders such as

CC cardiac arrest, cerebrovascular disorders such as cerebral ischaemia,
 CC nervous system disorders such as Alzheimer's disease, infections caused
 CC by bacteria, viruses and fungi, ocular disorders such as corneal
 CC infection, endocrine disorders such as premature labour and infertility,
 CC gastrointestinal disorders such as Crohn's disease, renal disorders such
 CC as glomerulonephritis and respiratory disorders such as asthma and
 CC pleurisy. The polypeptides can also be used to aid wound healing, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, to regenerate tissues and in chemotaxis.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

Query Match 1.7%; Score 18; DB 22; Length 489;
 Best Local Similarity 100.0%; Pred. No. 92;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 894 ACCCAAGCAGCCAGGACA 911
 DB 262 ACCCAAGCAGCCAGGACA 279
 |||||

RESULT 43
 AAV04274/c
 ID AAV04274 standard; cDNA; 733 BP.

XX AAV04274;

DT 22-JUN-1998 (first entry)

XX Secreted protein C195_4 cDNA.

XX Secreted protein; C195_4; peripheral blood mononuclear cell; PBMC;
 KW protein factor; human; ds.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT mat_peptide 13..733
 FT /*tag= a

PN W09748800-A1.

XX 24-DEC-1997.

XX 16-JUN-1997; 97WO-US10500.

XX 25-OCT-1996; 96US-0738367.

PR 17-JUN-1996; 96US-0664596.

PR 27-SEP-1996; 96US-0721926.

XX (GEMY) GENETICS INST INC.

XX Jacobs K, Lavallie ER, McCoy JM, Merberg D, Racie LA;

PI Spaulding V, Treacy M;

XX WPI; 1998-063141/06.

DR P-PSDB; AAW41601.

XX Nucleic acid encoding secreted protein from human peripheral blood

PT mononuclear cells - useful, e.g. as immunomodulators, antitumour

PT agents, promoters of tissue growth, haemostatic and thrombolytic

PT agents etc.

XX Claim 1; Page 37; 52pp; English.

XX This cDNA clone, designated C195_4, codes for a novel secreted

CC protein (see AAW41601) of human peripheral blood mononuclear cells

CC (PBMC). It was isolated from a PBMC cDNA library using methods

CC which are selective for cDNAs encoding secreted proteins. The 733

CC It is an additional isolate of clone C195_1 (see AAV04273), deposited
 CC as ATCC 98079. Both sequences can be recovered using a probe (see
 CC AAV04276). The isolated clone acid may be used to express recombinant
 CC secreted proteins; as a tissue/molecular weight marker; for
 CC chromosome identification; to identify possible genetic disorders;
 CC to isolate related DNA; as a source of PCR primers; to generate
 CC anti-protein or anti-DNA antibodies; in interaction trap assays to
 CC identify sequences that encode interacting proteins etc. The
 CC secreted proteins can be used to screen compounds for biological
 CC activity; to raise antibodies; as tissue markers; for isolation of
 CC related receptors and ligands and as nutritional sources. They may
 CC also have biological activities, e.g. cytokine, cell proliferation
 CC or differentiation activity, immunosuppressant, immunostimulant,
 CC regulation of haematopoiesis, modulation of fertility, chemotactic,
 CC chemokinetic, haemostatic, thrombolytic, antimicrobials, analgesic,
 CC antipsoriatic, etc. No evidence is given to support any of these
 CC activities or applications.

XX
 SQ Sequence 733 BP; 201 A; 152 C; 174 G; 206 T; 0 other;

Query Match 1.7%; Score 18; DB 19; Length 733;

Best Local Similarity 100.0%; Pred. No. 92;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 531 TGACATCATGTTCCAGCT 548

DB 341 TGACATCATGTTCCAGCT 324

RESULT 44
 AAI96369
 ID AAI96369 standard; cDNA; 757 BP.

XX AAI96369;

XX 13-NOV-2001 (first entry)

XX Human neuroblastoma expressed polynucleotide SEQ ID NO 2444.

XX Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.

XX Homo sapiens.

XX W0200166719-A1.

XX 13-SEP-2001.

XX 02-MAR-2001; 2001WO-JP01629.

XX 07-MAR-2000; 2000JP-0159195.

XX (CHIB-) CHIBA PREFECTURE.

XX (HISM) HISAMITSU PHARM CO LTD.

XX Nakagawara A;

XX WPI; 2001-565584/63.

XX Nucleic acids originating in gene expressed in human neuroblastoma,

PT useful as probe or primer in diagnosing prognosis of human

PT neuroblastoma, malignancy and susceptibility indicator or tumour marker

PT for anti-cancer agents -

XX Claim 1; Page 1798; 2979pp; Japanese.

XX The invention relates to novel genes (AAI93926-AAI97963) expressed in

CC human neuroblastoma. The nucleic acids are applicable as a probe or

CC primer in diagnosing the prognosis of human neuroblastoma, malignancy and

CC susceptibility indicators or tumour markers for anti-cancer agents. The

CC gene information for diagnosing prognosis is related to factors similar

CC to that for N-myc and TrkA genes.

XX Sequence 757 BP; 178 A; 147 C; 175 G; 248 T; 9 other;

Query Match 1.7%; Score 18; DB 22; Length 757;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 569 GCATCATCTATTATTTGCT 586
|||||
DB 350 GCATCATCTATTATTTGCT 367

RESULT 45

AA81592/c

ID AAA81592 standard; DNA; 767 BP.

XX AC

XX AAA81592;

XX XX

XX 04-DEC-2000 (first entry)

XX DE

XX N. meningitidis partial DNA sequence gnm_139 SEQ ID NO:139.

XX XX

XX Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;

KW antigen; vaccine; diagnosis; infection; antibacterial; identification;

KW Meningococcus B; MenB; ds.

XX XX

OS Neisseria meningitidis.

XX XX

XX WO200022430-A2.

XX PN

XX 20-APR-2000.

XX PD

XX 08-OCT-1999; 99WO-US23573.

XX PF

XX 09-OCT-1998; 98US-0103794.

XX PR

XX 30-APR-1999; 99US-0132068.

XX XX

XX (CHIR) CHIRON CORP.

XX PA

XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;

XX PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;

XX PI Rappuoli R, Pizza M;

XX XX

XX WPI; 2000-318079/27.

XX DR

XX XX

XX Isolated nucleotide sequences of Neisseria meningitidis which can be

XX PT used in the diagnosis and treatment of N. meningitidis infection and

XX PT other Neisserial infections, for example, N.gonorrhoea -

XX XX

XX Claim 7; Page 1508; 1760pp; English.

XX PS

XX The present invention describes methods of obtaining immunogenic

XX CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414

XX CC represent specifically claimed Neisseria meningitidis genomic DNA

XX CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent

XX CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to

XX CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the

XX CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to

XX CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF

XX CC sequences, which are all used in the exemplification of the present

XX CC invention. The nucleic acid sequences, protein sequences, and antibodies

XX CC against them, can be used in the manufacture of a composition. The

XX CC composition can be used as a medicament (or in the manufacture of a

XX CC medicament) for treating, preventing or diagnosing infection due to

XX CC Neisserial bacteria. For example, some of the identified proteins could

XX CC be components of vaccines against Meningococcus B; against all serotypes;

XX CC and/or against all pathogenic Neisseriae. Identification of sequences

XX CC from the bacterium will also facilitate production of biological probes,

XX CC particularly organism-specific probes. Attempts to make efficacious

XX CC Meningococcus B vaccines have failed mainly due to antigen tolerance.

XX CC Multivalent vaccines have also been tried but none have successfully

XX CC overcome antigenic variability. The provision of further, complete

XX CC sequences may provide an opportunity to identify secreted or surface

XX CC exposed proteins that may be presumed targets for the immune system and

XX CC which are not antigenically variable or at least more conserved than

CC other more variable regions.

XX XX

SQ Sequence 767 BP; 215 A; 144 C; 155 G; 253 T; 0 other;

Query Match 1.7%; Score 18; DB 21; Length 767;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 912 CTCAAAAACACAAAGGCC 929

|||||

DB 309 CTCAAAAACACAAAGGCC 292

|||||

Search completed: October 29, 2002, 03:48:33

Job time : 247 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2002, 03:17:32 ; Search time 1932 Seconds
(without alignments)
11275.634 Million cell updates/sec

Title: US-09-886-041-1
Perfect score: 1041
Sequence: 1 atgtacaacggctgctg.....acattgttgagtgcaactga 1041

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797636 seqs, 10463268293 residues

Word size : 0
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*	
1: gb_ba.*	AX147834 Sequence
2: gb_htg.*	AX148182 Sequence
3: gb_in.*	AX299705 Sequence
4: gb_om.*	AF411110 Homo sapi
5: gb_ov.*	AX338371 Sequence
6: gb_pat.*	AX338373 Sequence
7: gb_ph.*	AX305131 Sequence
8: gb_pl.*	AX277635 Sequence
9: gb_pr.*	AX299707 Sequence
10: gb_ro.*	AC026331 Homo sapi
11: gb_sts.*	AC026333 Homo sapi
12: gb_sy.*	AX338374 Sequence
13: gb_un.*	AX147766 Sequence
14: gb_vi.*	AX148194 Sequence
15: em_ba.*	AX335595 Sequence
16: em_fun.*	D10923 Human mRNA
17: em_hum.*	AX147888 Sequence
18: em_in.*	AX147917 Sequence
19: em_mu.*	AX147885 Sequence
20: em_om.*	AX147914 Sequence
21: em_or.*	AX147863 Sequence
22: em_ov.*	AX300198 Mus muscu
23: em_pat.*	AX300199 Mus muscu
24: em_ph.*	AL390918 Human DNA
25: em_pi.*	AC012158 Homo sapi
26: em_pl.*	AC011080 Homo sapi
27: em_pr.*	AC055724 Homo sapi
28: em_ro.*	AX147862 Sequence
29: em_sts.*	AB009589 Homo sapi
30: em_sy.*	AL161422 Human DNA
31: em_un.*	AL137848 Human DNA
32: em_vi.*	AP003771 Oryza sat
33: em_ba.*	AC006499 Homo sapi
34: em_fun.*	AC024933 Homo sapi
35: em_hum.*	AX147864 Sequence
36: em_in.*	AX147865 Sequence
37: em_mu.*	AX147886 Sequence
38: em_om.*	AX147915 Sequence
39: em_or.*	AX147887 Sequence
40: em_ov.*	AX147916 Sequence
41: em_pat.*	AB023806 Gallus ga
42: em_ph.*	AF224320 Gallus ga
43: em_pi.*	U88832 Human grom
44: em_ro.*	AC002997 Human DNA
45: em_sts.*	AC105178 Homo sapi

ALIGNMENTS

RESULT 1	AX147834	1041 bp	DNA	linear	PAT 08-JUN-2001
LOCUS	AX147834	Sequence 79 from Patent WO0136473.			
DEFINITION	AX147834	Novel g protein-coupled receptors			
ACCESSION	AX147834	Patent: WO 0136473-A 79 25-MAY-2001;			
VERSION	AX147834.1	PHARMACIA & UPJOHN COMPANY (US)			
KEYWORDS	human.	Location/Qualifiers			
SOURCE	human.	1..1041			
ORGANISM	Homo sapiens	/organism="Homo sapiens"			
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.				
	1 (bases 1 to 1041)				
	Vogeli,G., Wood,L.S., Parodi,L.A., Hiebsch,R.R., Lind,P.,				
	Slightom,J., Schellin,K.A., Kaytes,P.S., Bannigan,C.M., Ruff,V.,				
	Sejltz,T. and Huff,R.M.				
TITLE	Novel g protein-coupled receptors				
JOURNAL	Patent: WO 0136473-A 79 25-MAY-2001;				
FEATURES	PHARMACIA & UPJOHN COMPANY (US)				
source	Location/Qualifiers				
	1..1041				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
BASE COUNT	208 a	294 c	278 g	261 t	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description

ORIGIN

Query Match	100.0%;	Score 1041;	DB 6;	Length 1041;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1041;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGTACAAACGGTTCGTGCTGCGCATCAGGGGACACATCTCCAGTGTATGCGCGCG	60	
Db	1	ATGTACAAACGGTTCGTGCTGCGCATCAGGGGACACATCTCCAGTGTATGCGCGCG	60	
Qy	61	CTGCTCATTTGCGCTTTTGCTGGGCGCACTAGCGCAATGCGGTGCGCCTGTGTGGTTTC	120	
Db	61	CTGCTCATTTGCGCTTTTGCTGGGCGCACTAGCGCAATGCGGTGCGCCTGTGTGGTTTC	120	
Qy	121	TGCTTCCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTTCAATTTGGCCGTGGCT	180	
Db	121	TGCTTCCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTTCAATTTGGCCGTGGCT	180	
Qy	181	GATTTCCTCCTTATGATCTGCTGCTGCTTTTCGGACAGACTATTACCTCAGACGTAGACAC	240	
Db	181	GATTTCCTCCTTATGATCTGCTGCTTTTCGGACAGACTATTACCTCAGACGTAGACAC	240	
Qy	241	TGGGCTTTTGGGACATTCCTGCGGAGTGGGCTCTTCAGTTGGCCATGAACAGGCCC	300	
Db	241	TGGGCTTTTGGGACATTCCTGCGGAGTGGGCTCTTCAGTTGGCCATGAACAGGCCC	300	
Qy	301	GGGAGCATCGTGTTCCTTACGGTGGTGCCTGCGGACAGTATTTCAAAGTGGTCCACCCC	360	
Db	301	GGGAGCATCGTGTTCCTTACGGTGGTGCCTGCGGACAGTATTTCAAAGTGGTCCACCCC	360	
Qy	361	CACCACGGGTGAACACTATCTCCACCGGGTGGCGGTGGCATCGTCTGCACCTGTGG	420	
Db	361	CACCACGGGTGAACACTATCTCCACCGGGTGGCGGTGGCATCGTCTGCACCTGTGG	420	
Qy	421	GCCCTGGTCATCCTGGGAACAGTGTATCTTTTCTGGAGAACCACTCTGGTGCAGAG	480	
Db	421	GCCCTGGTCATCCTGGGAACAGTGTATCTTTTCTGGAGAACCACTCTGGTGCAGAG	480	
Qy	481	ACGGCCGTCTCCTGTGAGAGCTTCATCATGAGTGGCGCAATGGCTGGCATGACATCATG	540	
Db	481	ACGGCCGTCTCCTGTGAGAGCTTCATCATGAGTGGCGCAATGGCTGGCATGACATCATG	540	
Qy	541	TTCCAGCTGGAGTTCTTTATGCCCTCGGCATCATCTTATTTTGCTCCTTCAAGATTGTT	600	
Db	541	TTCCAGCTGGAGTTCTTTATGCCCTCGGCATCATCTTATTTTGCTCCTTCAAGATTGTT	600	
Qy	601	TGGACCTTGAGCGGAGGACGAGCTGCCAGACAGGCTCGGATCAAGAAGCGACCCGG	660	
Db	601	TGGACCTTGAGCGGAGGACGAGCTGCCAGACAGGCTCGGATCAAGAAGCGACCCGG	660	
Qy	661	TTTCATCATGGTGGCAATTTGTGTTTCATCATATGCTTACCTGCCAGCGTGTCTGCTAGA	720	
Db	661	TTTCATCATGGTGGCAATTTGTGTTTCATCATATGCTTACCTGCCAGCGTGTCTGCTAGA	720	
Qy	721	CTCTATTTCCTTGACCGGTCGCCCTCGAGTGGCTCGGATCCCTCTGTCATGGGCGCTG	780	
Db	721	CTCTATTTCCTTGACCGGTCGCCCTCGAGTGGCTCGGATCCCTCTGTCATGGGCGCTG	780	
Qy	781	CACATAACCCCTCAGCTTACACCTACATGAACAGCATGCTGGATCCCTGTGTATTATTTT	840	
Db	781	CACATAACCCCTCAGCTTACACCTACATGAACAGCATGCTGGATCCCTGTGTATTATTTT	840	
Qy	841	TCAAGCCCTCTGTTCCCAAATTCACAAACAGCTCAAAATCTGCAGTCTGAAACCCAG	900	
Db	841	TCAAGCCCTCTGTTCCCAAATTCACAAACAGCTCAAAATCTGCAGTCTGAAACCCAG	900	
Qy	901	CAGCCAGGACACTCAAAAACACAAAGGCGGAAGAGATGCCAAATTCGAACCTCGGTCCG	960	
Db	901	CAGCCAGGACACTCAAAAACACAAAGGCGGAAGAGATGCCAAATTCGAACCTCGGTCCG	960	
Qy	961	AGGAGTTCGATCAGTCTGGCAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC	1020	
Db	961	AGGAGTTCGATCAGTCTGGCAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC	1020	

Qy	1021	CACATTGTTGAGTGGCACTGA	1041	{ }
Db	1021	CACATTGTTGAGTGGCACTGA	1041	{ }
RESULT 2				
AX148182				
LOCUS				
DEFINITION Sequence 23 from Patent WO0136471.				
ACCESSION AX148182				
VERSION AX148182.1 GI:14347084				
KEYWORDS human.				
SOURCE Homo sapiens				
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE 1 (bases 1 to 1041)				
AUTHORS Chen,R., Dang,H.T. and Lowitz,K.P.				
TITLE Endogenous and non-endogenous versions of human g protein-coupled receptors				
JOURNAL Patent: WO 0136471-A 23 MAY-2001;				
ARENA Pharmaceuticals, Inc. (US)				
FEATURES Location/Qualifiers				
source 1..1041				
/organism="Homo sapiens"				
/db_xref="taxon:9606"				
BASE COUNT 208 a 294 c 278 g 261 t				
ORIGIN				
Query Match 100.0%; Score 1041; DB 6; Length 1041;				
Best Local Similarity 100.0%; Pred. No. 0;				
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1	ATGTACAACGGTGCGTCGCCGATCGAGGGGACACCATCTCCAGTGATCCGCCGC 60		
Db	1	ATGTACAACGGTGCGTCGCCGATCGAGGGGACACCATCTCCAGTGATCCGCCGC 60		
Qy	61	CTGCTCATTTGGCGCTTTGTGCTGGGCGACTAGGCAATGGGGTCGCCCTGTGTGTTTC 120		
Db	61	CTGCTCATTTGGCGCTTTGTGCTGGGCGACTAGGCAATGGGGTCGCCCTGTGTGTTTC 120		
Qy	121	TGCTTTCCACATGAAGACTTGAAGCCAGCAGCTGTTTACCTTTTCAATTTGGCCGTGGCT 180		
Db	121	TGCTTTCCACATGAAGACTTGAAGCCAGCAGCTGTTTACCTTTTCAATTTGGCCGTGGCT 180		
Qy	181	GATTTCCTCCTTATGATCTGCCTTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240		
Db	181	GATTTCCTCCTTATGATCTGCCTTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240		
Qy	241	TGGCGTTTTGGGGACATTCCTGCGAGTGGGGCTCTTACGTTGGCCATAAAGAGGCC 300		
Db	241	TGGCGTTTTGGGGACATTCCTGCGAGTGGGGCTCTTACGTTGGCCATAAAGAGGCC 300		
Qy	301	GGGAGCATCGTGTTCCTTACCGTGGTGGCTCGGACAGAGTATTTCAAAGTGGTCCACCCC 360		
Db	301	GGGAGCATCGTGTTCCTTACCGTGGTGGCTCGGACAGAGTATTTCAAAGTGGTCCACCCC 360		
Qy	361	CACCACGGGTGAACACTATATCCACCGGTGGCGGCTGGCATCGTCTGACCCCTGTGG 420		
Db	361	CACCACGGGTGAACACTATATCCACCGGTGGCGGCTGGCATCGTCTGACCCCTGTGG 420		
Qy	421	GCCTGTGTCATCTCGGAAACAGTGTATCTTTTGTGGAGAACCATTCTCTGGTGCACAG 480		
Db	421	GCCTGTGTCATCTCGGAAACAGTGTATCTTTTGTGGAGAACCATTCTCTGGTGCACAG 480		
Qy	481	ACGGCCGTCTCCTGTGAGAGCTTCATCATGAGGTGGGCAATGGCTGGCATGACATCATG 540		
Db	481	ACGGCCGTCTCCTGTGAGAGCTTCATCATGAGGTGGGCAATGGCTGGCATGACATCATG 540		
Qy	541	TTCAGCTGGAGTTCCTTTATGCCCTCGGCATCATCTTATTTTGGTCTCCTCAAGATTGTT 600		
Db	541	TTCAGCTGGAGTTCCTTTATGCCCTCGGCATCATCTTATTTTGGTCTCCTCAAGATTGTT 600		

[illegible]


```
QY 1 ATGTACAACGGGTCTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 60
Db 6 ATGTACAACGGGTCTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 65
QY 61 CTGCTCATTTGTGGCCCTTTGTGCTGGCGGCACTAGGCAATGGGGTCCGCTGTGTGTTTC 120
Db 66 CTGCTCATTTGTGGCCCTTTGTGCTGGCGGCACTAGGCAATGGGGTCCGCTGTGTGTTTC 125
QY 121 TGTCTCCACATGAAGACCTGGAGGCCAGCACGTGTTTACCTTTTCAATTTGGCCGTGGCT 180
Db 126 TGTCTCCACATGAAGACCTGGAGGCCAGCACGTGTTTACCTTTTCAATTTGGCCGTGGCT 185
QY 181 GATTTCCTCTTATGATCTGCTGCTCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
Db 186 GATTTCCTCTTATGATCTGCTGCTCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 245
QY 241 TGGGCTTTTGGGACATTCCTCCGAGTGGGCTCTTACGTTGGCCATGAACAGGCC 300
Db 246 TGGGCTTTTGGGACATTCCTCCGAGTGGGCTCTTACGTTGGCCATGAACAGGCC 305
QY 301 GGGAGCATCGTGTCTTACGGTGGTGGCTGGCGACAGGTATTTCAAAGTGGTCCACCCC 360
Db 306 GGGAGCATCGTGTCTTACGGTGGTGGCTGGCGACAGGTATTTCAAAGTGGTCCACCCC 365
QY 361 CACCACGCGGTGAACACTATCTCCACCGGGTGGCGCTGGCATCTCTGCACCCCTGTG 420
Db 366 CACCACGCGGTGAACACTATCTCCACCGGGTGGCGCTGGCATCTCTGCACCCCTGTG 425
QY 421 GCCCTGGTCTCTCTGGACAGTGTATCTTTTCTGCTGGAAACCATCTCTGCGTGAAGAG 480
Db 426 GCCCTGGTCTCTCTGGACAGTGTATCTTTTCTGCTGGAAACCATCTCTGCGTGAAGAG 485
QY 481 ACGGCGCTCTCTCTGAGAGCTTCATCATGGAGTGGCGCAATGGCTGGCATGACATCATG 540
Db 486 ACGGCGCTCTCTCTGAGAGCTTCATCATGGAGTGGCGCAATGGCTGGCATGACATCATG 545
QY 541 TTCCAGCTGGAGTCTCTTTATGCCCTCGGCATCATCTTATTTTGCCTCTCAAGATTGT 600
Db 546 TTCCAGCTGGAGTCTCTTTATGCCCTCGGCATCATCTTATTTTGCCTCTCAAGATTGT 605
QY 601 TGGAGCTTGAGGCGGAGGACAGCTGGCGAGACAGCTCGGATGAAGAAGGCGACCCCG 660
Db 606 TGGAGCTTGAGGCGGAGGACAGCTGGCGAGACAGCTCGGATGAAGAAGGCGACCCCG 665
QY 661 TTCATCATGCTGTGCAATTTGTTTCATCAGATGCTACCTGCCAGCGTGTCTGCTAGA 720
Db 666 TTCATCATGCTGTGCAATTTGTTTCATCAGATGCTACCTGCCAGCGTGTCTGCTAGA 725
QY 721 CTCATTTCTCTGGAGTGGCCCTCGAGTGGCTGGCATCCCTCTGTCCATGGGGCCCTG 780
Db 726 CTCATTTCTCTGGAGTGGCCCTCGAGTGGCTGGCATCCCTCTGTCCATGGGGCCCTG 785
QY 781 CACATTAACCTCAGTTCACCTACATGAACAGCATGCTGGATCCCTCGTGTATATTTT 840
Db 786 CACATTAACCTCAGTTCACCTACATGAACAGCATGCTGGATCCCTCGTGTATATTTT 845
QY 841 TCAAGCCCTCTCTTCCCAATTTCTCAACAAGCTCAAAATCTGCAGTCTGAAACCCCAAG 900
Db 846 TCAAGCCCTCTCTTCCCAATTTCTCAACAAGCTCAAAATCTGCAGTCTGAAACCCCAAG 905
QY 901 CAGCCAGGACACTCAAAAACACAAAGCCGGAAGAGATGCCAAATTCGAACTCGGTCCG 960
Db 906 CAGCCAGGACACTCAAAAACACAAAGCCGGAAGAGATGCCAAATTCGAACTCGGTCCG 965
QY 961 AGGAGTTGCATAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGCAATGGATCCC 1020
Db 966 AGGAGTTGCATAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGCAATGGATCCC 1025
QY 1021 CACATTGTTGAGTGGCACTGA 1041
Db 1026 CACATTGTTGAGTGGCACTGA 1046
```

```
RESULT 6
AX338373
LOCUS AX338373 1050 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 3 from Patent WO0174904.
ACCESSION AX338373
VERSION AX338373.1 GI:18128870
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (sites)
AUTHORS Majumder,K., Vernet,C.A., Casman,S.J., Wolenc,A.R., Spaderna,S.K.,
Padigaru,M., Mishnu,V.S., Tchernev,V.T., Spytek,K.A., Li,L.,
Baumgartner,J.C. and Gusev,V.Y.
TITLE Novel proteins and nucleic acids encoding same
JOURNAL Patent: WO 0174904-A 3 11-OCT-2001;
FEATURES Location/Qualifiers
source 1..1050
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 211 a 298 c 279 g 262 t
ORIGIN
Query Match 100.0%; Score 1041; DB 6; Length 1050;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTACAACGGGTCTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 60
Db 6 ATGTACAACGGGTCTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 65
QY 61 CTGCTCATTTGTGGCCCTTTGTGCTGGCGGCACTAGGCAATGGGGTCCGCTGTGTGTTTC 120
Db 66 CTGCTCATTTGTGGCCCTTTGTGCTGGCGGCACTAGGCAATGGGGTCCGCTGTGTGTTTC 125
QY 121 TGTCTCCACATGAAGACCTGGAGGCCAGCACGTGTTTACCTTTTCAATTTGGCCGTGGCT 180
Db 126 TGTCTCCACATGAAGACCTGGAGGCCAGCACGTGTTTACCTTTTCAATTTGGCCGTGGCT 185
QY 181 GATTTCCTCTTATGATCTGCTGCTCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
Db 186 GATTTCCTCTTATGATCTGCTGCTCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 245
QY 241 TGGGCTTTTGGGACATTCCTCCGAGTGGGCTCTTACGTTGGCCATGAACAGGCC 300
Db 246 TGGGCTTTTGGGACATTCCTCCGAGTGGGCTCTTACGTTGGCCATGAACAGGCC 305
QY 301 GGGAGCATCGTGTCTTACGGTGGTGGCTGGCGACAGGTATTTCAAAGTGGTCCACCCC 360
Db 306 GGGAGCATCGTGTCTTACGGTGGTGGCTGGCGACAGGTATTTCAAAGTGGTCCACCCC 365
QY 361 CACCACGCGGTGAACACTATCTCCACCGGGTGGCGCTGGCATCTCTGCACCCCTGTG 420
Db 366 CACCACGCGGTGAACACTATCTCCACCGGGTGGCGCTGGCATCTCTGCACCCCTGTG 425
QY 421 GCCCTGGTCTCTCTGGACAGTGTATCTTTTCTGCTGGAAACCATCTCTGCGTGAAGAG 480
Db 426 GCCCTGGTCTCTCTGGACAGTGTATCTTTTCTGCTGGAAACCATCTCTGCGTGAAGAG 485
QY 481 ACGGCGCTCTCTCTGAGAGCTTCATCATGGAGTGGCGCAATGGCTGGCATGACATCATG 540
Db 486 ACGGCGCTCTCTCTGAGAGCTTCATCATGGAGTGGCGCAATGGCTGGCATGACATCATG 545
QY 541 TTCCAGCTGGAGTCTCTTTATGCCCTCGGCATCATCTTATTTTGCCTCTCAAGATTGT 600
Db 546 TTCCAGCTGGAGTCTCTTTATGCCCTCGGCATCATCTTATTTTGCCTCTCAAGATTGT 605
QY 601 TGGAGCTTGAGGCGGAGGACAGCTGGCGAGACAGCTCGGATGAAGAAGGCGACCCCG 660
Db 606 TGGAGCTTGAGGCGGAGGACAGCTGGCGAGACAGCTCGGATGAAGAAGGCGACCCCG 665
QY 661 TTCATCATGCTGTGCAATTTGTTTCATCAGATGCTACCTGCCAGCGTGTCTGCTAGA 720
Db 666 TTCATCATGCTGTGCAATTTGTTTCATCAGATGCTACCTGCCAGCGTGTCTGCTAGA 725
QY 721 CTCATTTCTCTGGAGTGGCCCTCGAGTGGCTGGCATCCCTCTGTCCATGGGGCCCTG 780
Db 726 CTCATTTCTCTGGAGTGGCCCTCGAGTGGCTGGCATCCCTCTGTCCATGGGGCCCTG 785
QY 781 CACATTAACCTCAGTTCACCTACATGAACAGCATGCTGGATCCCTCGTGTATATTTT 840
Db 786 CACATTAACCTCAGTTCACCTACATGAACAGCATGCTGGATCCCTCGTGTATATTTT 845
QY 841 TCAAGCCCTCTCTTCCCAATTTCTCAACAAGCTCAAAATCTGCAGTCTGAAACCCCAAG 900
Db 846 TCAAGCCCTCTCTTCCCAATTTCTCAACAAGCTCAAAATCTGCAGTCTGAAACCCCAAG 905
QY 901 CAGCCAGGACACTCAAAAACACAAAGCCGGAAGAGATGCCAAATTCGAACTCGGTCCG 960
Db 906 CAGCCAGGACACTCAAAAACACAAAGCCGGAAGAGATGCCAAATTCGAACTCGGTCCG 965
QY 961 AGGAGTTGCATAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGCAATGGATCCC 1020
Db 966 AGGAGTTGCATAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGCAATGGATCCC 1025
QY 1021 CACATTGTTGAGTGGCACTGA 1041
Db 1026 CACATTGTTGAGTGGCACTGA 1046
```

```
QY 661 TTATCATATGGTGGGCAATTGTGTTTCATCAGATGCTACCTGCCAGCGTGCTGCTAGA 720
|||||
Db 666 TTATCATATGGTGGGCAATTGTGTTTCATCAGATGCTACCTGCCAGCGTGCTGCTAGA 725
|||||
QY 721 CTCTATTTCCTCTGGAGCGGTGCCCTCAGTGCCTGCCATCCCTCTGTCATGGGCGCTG 780
|||||
Db 726 CTCTATTTCCTCTGGAGCGGTGCCCTCAGTGCCTGCCATCCCTCTGTCATGGGCGCTG 785
|||||
QY 781 CACATAACCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTGGTGATATTTT 840
|||||
Db 786 CACATAACCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTGGTGATATTTT 845
|||||
QY 841 TCAAGCCCTCTCTTCCCAATTTCTACAACAAGCTCAAAATCTGAGTCTGAAACCCAAAG 900
|||||
Db 846 TCAAGCCCTCTCTTCCCAATTTCTACAACAAGCTCAAAATCTGAGTCTGAAACCCAAAG 905
|||||
QY 901 CAGCCAGGACACTCAAAAACACAAAGCCGGAAGAGATGCCAATTTTCGAACCTCGGTGCG 960
|||||
Db 906 CAGCCAGGACACTCAAAAACACAAAGCCGGAAGAGATGCCAATTTTCGAACCTCGGTGCG 965
|||||
QY 961 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGCAATGGGATCCC 1020
|||||
Db 966 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGCAATGGGATCCC 1025
|||||
QY 1021 CACATTGTTGAGTGGCACTGA 1041
|||||
Db 1026 CACATTGTTGAGTGGCACTGA 1046
|||||

RESULT 7
AX305131 1083 bp DNA linear PAT 11-DEC-2001
LOCUS
DEFINITION Sequence 11 from Patent WO0187937.
ACCESSION AX305131
VERSION AX305131.1 GI:17644766
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (sites)
Patterson,C., Lu,D.A., Thornton,M., Lu,Y., Tribouley,C.M.,
Graul,R., Khan,F.A., Gandhi,A.R., Walla,N.K., Nguyen,D.B., Yue,H.,
Hafalia,A., Elliott,V.S., Lal,P., Reddy,R., Kallick,D.A., Tang,T.Y.
and Au-Young,J.
G-protein coupled receptors
Patent: WO 0187937-A 11 22-NOV-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
1..1083
/organism="Homo sapiens"
/db_xref="taxon:9606"
/notes="Incyte ID No: 7474846CB1"

BASE COUNT 211 a 314 c 288 g 270 t
ORIGIN

Query Match 100.0%; Score 1041; DB 6; Length 1083;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTACACGGGTGCTGCTGCCGATCGAGGGGACACCATCTCCAGAGTGATGCCGCG 60
|||||
Db 43 ATGTACACGGGTGCTGCTGCCGATCGAGGGGACACCATCTCCAGAGTGATGCCGCG 102
|||||
QY 61 CTGCTCATTTGGCGCTTTGCTGGCGCACTAGGCAATGGGTCGCCCTGTGTTTC 120
|||||
Db 103 CTGCTCATTTGGCGCTTTGCTGGCGCACTAGGCAATGGGTCGCCCTGTGTTTC 162
|||||
QY 121 TGCTTCCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTTCAATTTGGCGGTGCT 180
|||||
Db 163 TGCTTCCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTTCAATTTGGCGGTGCT 222
|||||
QY 181 GATTTCCTCTTATGATCTGCCTGCCTTTTCGGAGACAGACTATTACCTCAGACGTAGACAC 240
|||||
```

```
Db 223 GATTTCCTCTTATGATCTGCCTGCCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 282
|||||
QY 241 TGGGCTTTTGGGACATTCCTCTGCCAGTGGGCTCTTACAGTTGGCCATGAACAGGGCC 300
|||||
Db 283 TGGGCTTTTGGGACATTCCTCTGCCAGTGGGCTCTTACAGTTGGCCATGAACAGGGCC 342
|||||
QY 301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGATTTTCAAAGTGGTCCACCC 360
|||||
Db 343 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGATTTTCAAAGTGGTCCACCC 402
|||||
QY 361 CACACCGGTGAACACTATCTCCACCCGGGTGGCGGCTGGCATCGTCTGCACCCCTGTGG 420
|||||
Db 403 CACACCGGTGAACACTATCTCCACCCGGGTGGCGGCTGGCATCGTCTGCACCCCTGTGG 462
|||||
QY 421 GCCTGTGTCATCTCTGGGAACAGTGTATCTTTTGTCTGAGAACCATCTCTCGGTGCAAGAG 480
|||||
Db 463 GCCTGTGTCATCTCTGGGAACAGTGTATCTTTTGTCTGAGAACCATCTCTCGGTGCAAGAG 522
|||||
QY 481 ACGGCCGTCTCTGTGAGAGCTTTCATCATGAGTTCGCGCAATGGCTGGCATGACATCATG 540
|||||
Db 523 ACGGCCGTCTCTGTGAGAGCTTTCATCATGAGTTCGCGCAATGGCTGGCATGACATCATG 582
|||||
QY 541 TTCAGCTGGAGTCTTTATGCCCCCTCGGCATCATCTTATTTTCTCTCCTTCAAGATTGTT 600
|||||
Db 583 TTCAGCTGGAGTCTTTATGCCCCCTCGGCATCATCTTATTTTCTCTCCTTCAAGATTGTT 642
|||||
QY 601 TGGAGCCTGAGGGGAGGAGCAGCAGCTGCCAGACAGAGCTCGGATGAAGAGGCGACCCGG 660
|||||
Db 643 TGGAGCCTGAGGGGAGGAGCAGCAGCTGCCAGACAGAGCTCGGATGAAGAGGCGACCCGG 702
|||||
QY 661 TTCAATCATGTGTGGCAATGTTGTTTCATCAGATGCTACCTGCCAGCGTGCTGCTAGA 720
|||||
Db 703 TTCATCATGTGTGTGGCAATGTTGTTTCATCAGATGCTACCTGCCAGCGTGCTGCTAGA 762
|||||
QY 721 CTCTATTTCCTCTGGAGCGGTGCCCTCGAGTGCCTGCCATCCCTCTGTCATGGGCGCTG 780
|||||
Db 763 CTCTATTTCCTCTGGAGCGGTGCCCTCGAGTGCCTGCCATCCCTCTGTCATGGGCGCTG 822
|||||
QY 781 CACATAACCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTGGTGATATTTT 840
|||||
Db 823 CACATAACCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTGGTGATATTTT 882
|||||
QY 841 TCAAGCCCTCTCTTCCCAATTTCTACAACAAGCTCAAAATCTGCACTCTGAAACCCAAAG 900
|||||
Db 883 TCAAGCCCTCTCTTCCCAATTTCTACAACAAGCTCAAAATCTGCACTCTGAAACCCAAAG 942
|||||
QY 901 CAGCCAGGACACTCAAAAACACAAAGCCGGAAGAGATGCCAATTTTCGAACCTCGGTGCG 960
|||||
Db 943 CAGCCAGGACACTCAAAAACACAAAGCCGGAAGAGATGCCAATTTTCGAACCTCGGTGCG 1002
|||||
QY 961 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020
|||||
Db 1003 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1062
|||||
QY 1021 CACATTGTTGAGTGGCACTGA 1041
|||||
Db 1063 CACATTGTTGAGTGGCACTGA 1083
|||||

RESULT 8
AX277635 1730 bp DNA linear PAT 01-NOV-2001
LOCUS
DEFINITION Sequence 1 from Patent WO0177320.
ACCESSION AX277635
VERSION AX277635.1 GI:16604811
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (sites)
Xiao,Y.
REFERENCE
AUTHORS
```

TITLE Regulation of human hm74-like g protein coupled receptor
JOURNAL Patent: WO 0177320-A 1 18-OCT-2001;
Bayer Aktiengesellschaft (DE)

FEATURES
source Location/Qualifiers

BASE COUNT 361 a 494 c 444 g 431 t
ORIGIN

Query Match 100.0%; Score 1041; DB 6; Length 1730;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTACAAGGGTCGTCGCCCATCGAGGGGACACCATCTCCAGGTGATGCCGCG 60
|||||
DB 464 ATGTACAAGGGTCGTCGCCCATCGAGGGGACACCATCTCCAGGTGATGCCGCG 523
QY 61 CTGCTCATTTGTGGCCCTTTGTCTGGCGCAGTGGCAATGGGTCGCCCTGTGTGTTTC 120
|||||
DB 524 CTGCTCATTTGTGGCCCTTTGTCTGGCGCAGTGGCAATGGGTCGCCCTGTGTGTTTC 583
QY 121 TGCCTCCACATGAAGACCTTGAAGCCCGCAGCTGTTTACCTTTCAATTTGGCCGTGGCT 180
|||||
DB 584 TGCCTCCACATGAAGACCTTGAAGCCCGCAGCTGTTTACCTTTCAATTTGGCCGTGGCT 643
QY 181 GATTTCCTCTTATGATCTGCCCTCTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
|||||
DB 644 GATTTCCTCTTATGATCTGCCCTCTTTCGGACAGACTATTACCTCAGACGTAGACAC 703
QY 241 TGGGCTTTTGGGACATTCCTCCGAGTGGGCTCTTCACGTTGGCCCATGAACAGGGCC 300
|||||
DB 704 TGGGCTTTTGGGACATTCCTCCGAGTGGGCTCTTCACGTTGGCCCATGAACAGGGCC 763
QY 301 GGGAGCATGCTTCTTACGCTGGTGGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC 360
|||||
DB 764 GGGAGCATGCTTCTTACGCTGGTGGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC 823
QY 361 CACCACGGGTGAACACTATCTCACCCCGGTGGCGCTGGCATCTCTGCACCCCTGTGG 420
|||||
DB 824 CACCACGGGTGAACACTATCTCACCCCGGTGGCGCTGGCATCTCTGCACCCCTGTGG 883
QY 421 GCCTGGTCATCTCGGGAACAGTGTATCTTTTGTCTGGGAACCATCTCTCGCTGCAAGAG 480
|||||
DB 884 GCCTGGTCATCTCGGGAACAGTGTATCTTTTGTCTGGGAACCATCTCTCGCTGCAAGAG 943
QY 481 ACGGCGCTCTCTCTGAGAGCTTCATCAGGAGTGGGCAATGGCTGGCATGACATCATG 540
|||||
DB 944 ACGGCGCTCTCTCTGAGAGCTTCATCAGGAGTGGGCAATGGCTGGCATGACATCATG 1003
QY 541 TTCACGCTGGAGTCTTTATGCCCTCGGCATCATCTTATTTGCTTCCCTCAAGATTGTT 600
|||||
DB 1004 TTCACGCTGGAGTCTTTATGCCCTCGGCATCATCTTATTTGCTTCCCTCAAGATTGTT 1063
QY 601 TGGAGCTGAGGCGGAGCAGCTGGCCAGACAGGCTCGGATGAAGAAGGCGACCCCG 660
|||||
DB 1064 TGGAGCTGAGGCGGAGCAGCTGGCCAGACAGGCTCGGATGAAGAAGGCGACCCCG 1123
QY 661 TTCATCATGTTGGCAATTGTGTTTCATCATCATGCTACCTGCCCCAGCGTGTCTGTAGA 720
|||||
DB 1124 TTCATCATGTTGGCAATTGTGTTTCATCATCATGCTACCTGCCCCAGCGTGTCTGTAGA 1183
QY 721 CTCTATTCTCTGAGGCGGCTCGAGTGGCTGGATCCCTCTCTCCATGGGGCCCTG 780
|||||
DB 1184 CTCTATTCTCTGAGGCGGCTCGAGTGGCTGGATCCCTCTCTCCATGGGGCCCTG 1243
QY 781 CACATAACCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTGGTGTATTATTTT 840
|||||
DB 1244 CACATAACCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTGGTGTATTATTTT 1303
QY 841 TCAAGCCCTCTCTTCCCAAAATTTCTACAACAGCTCAAAATCTGCAGTCTGAACCCCAAG 900
|||||
DB 1304 TCAAGCCCTCTCTTCCCAAAATTTCTACAACAGCTCAAAATCTGCAGTCTGAACCCCAAG 1363

QY 901 CAGCCAGGACACTCAAAAACACAAAGCCGGAAGAGATGCCAATTTTGAACCTCGGTGCG 960
|||||
DB 1364 CAGCCAGGACACTCAAAAACACAAAGCCGGAAGAGATGCCAATTTTGAACCTCGGTGCG 1423
QY 961 AGGAGTTGCATCAGTGTGGCAATAGTTCACAAAGCCAGTCTGTATGGGAATGGGATCCC 1020
|||||
DB 1424 AGGAGTTGCATCAGTGTGGCAATAGTTCACAAAGCCAGTCTGTATGGGAATGGGATCCC 1483
QY 1021 CACATTGTTGAGTGGCACTGA 1041
|||||
DB 1484 CACATTGTTGAGTGGCACTGA 1504
RESULT 9
AX299707
LOCUS AX299707 2331 bp DNA linear PAT 26-NOV-2001
DEFINITION Sequence 3 from Patent WO0173029.
ACCESSION AX299707
VERSION AX299707.1 GI:17129251
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Ye,J.C., Cravchik,A.C., di Francesco,V.C. and Beasley,E.M.
TITLE Isolated human g-protein coupled receptors, nucleic acid molecules
encoding human gpcr proteins, and uses thereof
JOURNAL Patent: WO 0173029-A 3 04-OCT-2001;
PE Corporation (NV) (US)
FEATURES
source Location/Qualifiers
1..2331
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 497 a 620 c 592 g 622 t
ORIGIN
Query Match 100.0%; Score 1041; DB 6; Length 2331;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTACAAGGGTCGTCGCCCATCGAGGGGACACCATCTCCAGGTGATGCCGCGCG 60
|||||
DB 311 ATGTACAAGGGTCGTCGCCCATCGAGGGGACACCATCTCCAGGTGATGCCGCGCG 370
QY 61 CTGCTCATTTGTGGCCCTTTGTCTGGCGCAGTGGCAATGGGTCGCCCTGTGTGTTTC 120
|||||
DB 371 CTGCTCATTTGTGGCCCTTTGTCTGGCGCAGTGGCAATGGGTCGCCCTGTGTGTTTC 430
QY 121 TGCCTCCACATGAAGACCTTGAAGCCCGCAGCTGTTTACCTTTTCAATTTGGCCGTGGCT 180
|||||
DB 431 TGCCTCCACATGAAGACCTTGAAGCCCGCAGCTGTTTACCTTTTCAATTTGGCCGTGGCT 490
QY 181 GATTTCCTCTCTTATGATCTGCCCTCTTTCGGACAGACTATTACCTCAGACCTAGACAC 240
|||||
DB 491 GATTTCCTCTCTTATGATCTGCCCTCTTTCGGACAGACTATTACCTCAGACCTAGACAC 550
QY 241 TGGGCTTTTGGGACATTCCTCCGAGTGGGCTCTTCACGTTGGCCATGAACAGGGCC 300
|||||
DB 551 TGGGCTTTTGGGACATTCCTCCGAGTGGGCTCTTCACGTTGGCCATGAACAGGGCC 610
QY 301 GGGAGCATGCTGTTCTTACGTTGGTGGCTGGCGACAGGTATTTTCAAGTGGTCCACCCC 360
|||||
DB 611 GGGAGCATGCTGTTCTTACGTTGGTGGCTGGCGACAGGTATTTTCAAGTGGTCCACCCC 670
QY 361 CACCACGGGTGAACACTATCTCACCCCGGTGGCGCTGGCATCGTCTGCACCCCTGTGG 420
|||||
DB 671 CACCACGGGTGAACACTATCTCACCCCGGTGGCGCTGGCATCGTCTGCACCCCTGTGG 730
QY 421 GCCTGGTCATCTCTGGGAACAGTGTATCTTTTGTCTGGGAACCATCTCTCGCTGCAAGAG 480
|||||
DB 731 GCCTGGTCATCTCTGGGAACAGTGTATCTTTTGTCTGGGAACCATCTCTCGCTGCAAGAG 790

```
Qy 481 ACGGCGCTCTCTGTGAGAGCTTTCATCATGAGTCGGCGCAATGCTGGCATGACATCATG 540
|||||
Db 791 ACGGCGCTCTCTGTGAGAGCTTTCATCATGAGTCGGCGCAATGCTGGCATGACATCATG 850
|||||
Qy 541 TTCAGCTGGAGTCTTTATGCCCCCTCGGCATCATCTTATTTTCTCTCTCTCAAGATTGTT 600
|||||
Db 851 TTCAGCTGGAGTCTTTATGCCCCCTCGGCATCATCTTATTTTCTCTCTCTCAAGATTGTT 910
|||||
Qy 601 TGGAGCTGAGGCGGAGCGACGCTGCCAGACAGCTCGGATGAAGAGCGCCACCCGG 660
|||||
Db 911 TGGAGCTGAGGCGGAGCGACGCTGCCAGACAGCTCGGATGAAGAGCGCCACCCGG 970
|||||
Qy 661 TTCATCATGTGTGGCAATTTGTTTCATCATGCTACCTGCCAGCGTCTCTGCTAGA 720
|||||
Db 971 TTCATCATGTGTGGCAATTTGTTTCATCATGCTACCTGCCAGCGTCTCTGCTAGA 1030
|||||
Qy 721 CTCATTTCCTCTGGAGCGTCCCTCGAGTGCCTGCGATGCCCTCTGTCATGCGGCCCTG 780
|||||
Db 1031 CTCATTTCCTCTGGAGCGTCCCTCGAGTGCCTGCGATGCCCTCTGTCATGCGGCCCTG 1090
|||||
Qy 781 CACATAACCTCAGCTTCACCTACATGACAGCATGCTGGATGCCCTGCTATTATTTT 840
|||||
Db 1091 CACATAACCTCAGCTTCACCTACATGACAGCATGCTGGATGCCCTGCTATTATTTT 1150
|||||
Qy 841 TCAAGCCCTCTCTTCCCAATTTAGCAACAGCTCAAAATCTGCAGTCTGAAACCCCAAG 900
|||||
Db 1151 TCAAGCCCTCTCTTCCCAATTTAGCAACAGCTCAAAATCTGCAGTCTGAAACCCCAAG 1210
|||||
Qy 901 CAGCAGGACACTCAAAACACAAAGCGGAGAGATGCCAATTTGCAACCTCGGTGCG 960
|||||
Db 1211 CAGCAGGACACTCAAAACACAAAGCGGAGAGATGCCAATTTGCAACCTCGGTGCG 1270
|||||
Qy 961 AGGAGTTCATCAGCTGGCAATAGTTTCCAAAGCAGTCTGTATGGCAATGGATGCC 1020
|||||
Db 1271 AGGAGTTCATCAGCTGGCAATAGTTTCCAAAGCAGTCTGTATGGCAATGGATGCC 1330
|||||
Qy 1021 CACATTGTTGAGTGGCACTGA 1041
|||||
Db 1331 CACATTGTTGAGTGGCACTGA 1351
|||||
RESULT 10
AC026331
LOCUS Homo sapiens chromosome 12 clone RP11-507N20, WORKING DRAFT
DEFINITION
AC026331
ACCESSION
VERSION AC026331.18 GI:14547388
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 179172)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Duggan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jollivet,S.,
```

Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Liu,C., Liu,J., Liu,W.,
Loulsegah,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S.,
Ogah,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pichens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,
Ruiz,S., Saverly,G., Scherer,S., Scott,G., Shen,H., Shoshitari,N.,
Slisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 179172)

Worley, K.C.

Direct Submission

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (22-MAR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jun 25, 2001 this sequence version replaced gi:13877178.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: HAQI

Center clone name: RP11-507N20

----- Summary Statistics

Sequencing vector: Plasmid; M7789

Sequencing vector: M13; L08821

Chemistry: Dye-primer Bodipy; 31% of reads

Chemistry: Dye-terminator Big Dye; 69% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 175811 bases at least Q40

Consensus quality: 178545 bases at least Q30

Consensus quality: 179678 bases at least Q20

Estimated insert size: 176741; sum-of-contigs estimation

Quality coverage: 0x in Q20 bases; agarose-gel estimation

Quality coverage: 9.4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 10 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 78354: contig of 78354 bp in length

* 78355 78454: gap of unknown length

* 78455 104278: contig of 25824 bp in length

* 104279 104378: gap of unknown length

* 104379 125803: contig of 21425 bp in length

* 125804 125903: gap of unknown length

* 125904 145254: contig of 19351 bp in length

* 145255 145354: gap of unknown length

* 145355 153279: contig of 7925 bp in length

* 153280 153379: gap of unknown length

* 153380 160977: contig of 7598 bp in length

* 160978 161077: gap of unknown length

* 161078 167351: contig of 6274 bp in length
* 167352 167451: gap of unknown length
* 167452 173389: contig of 5938 bp in length
* 173390 173489: gap of unknown length
* 173490 176258: contig of 2769 bp in length
* 176259 176358: gap of unknown length
* 176359 179172: contig of 2814 bp in length.

FEATURES
Location/Qualifiers
source
1..179172
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-507N20"
BASE COUNT 49256 a 39904 c 41184 g 47913 t 915 others
ORIGIN

Query Match		100.0%;	Score 1041;	DB 2;	Length 179172;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 1041;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGTACAAGGGTCGTGCTGCCCATCGAGGGGACACCATCTCCAGGTGATGCGCGG	60		
DB	11346	ATGTACAAGGGTCGTGCTGCCCATCGAGGGGACACCATCTCCAGGTGATGCGCGG	11405		
QY	61	CTGCTCATTTGGCCCTTTGTGCTGGCGCACTAGGCAATGGGTCGCCCTGTGTGTTTC	120		
DB	11406	CTGCTCATTTGGCCCTTTGTGCTGGCGCACTAGGCAATGGGTCGCCCTGTGTGTTTC	11465		
QY	121	TGCTTCACATGAAGACCTGGGAAGCCAGACATGTTTACCTTTTCAATTTGGCCGTGGCT	180		
DB	11466	TGCTTCACATGAAGACCTGGGAAGCCAGACATGTTTACCTTTTCAATTTGGCCGTGGCT	11525		
QY	181	GATTTCTCTCTTATGATCGCTCGCTTTTCGGACAGACTATTACCTCAGACGTACAC	240		
DB	11526	GATTTCTCTCTTATGATCGCTCGCTTTTCGGACAGACTATTACCTCAGACGTACAC	11585		
QY	241	TGGGCTTTTGGGACATTTCCCTGCCAGTGGGCTCTTCACGTTGGCCATGAACAGGGC	300		
DB	11586	TGGGCTTTTGGGACATTTCCCTGCCAGTGGGCTCTTCACGTTGGCCATGAACAGGGC	11645		
QY	301	GGGAGCATGTTCTTACGGTGGTGGCTGGCGAGAGGTATTTCAAGTGGTCCACCCC	360		
DB	11646	GGGAGCATGTTCTTACGGTGGTGGCTGGCGAGAGGTATTTCAAGTGGTCCACCCC	11705		
QY	361	CACACGCGGTGAACACTATCTCCACCGGTGGCGGTGGCATCGTCTGCACCCCTGTGG	420		
DB	11706	CACACGCGGTGAACACTATCTCCACCGGTGGCGGTGGCATCGTCTGCACCCCTGTGG	11765		
QY	421	GCCTGGTCACTCTGGGAACAGTGTATCTTTTCTGGAGAACCATCTCTGCGTGCAGAG	480		
DB	11766	GCCTGGTCACTCTGGGAACAGTGTATCTTTTCTGGAGAACCATCTCTGCGTGCAGAG	11825		
QY	481	ACGGCGCTCTCTGTAGAGCTTCATCATGGAGTGGCCCAATGGCTGGCATGACATCATG	540		
DB	11826	ACGGCGCTCTCTGTAGAGCTTCATCATGGAGTGGCCCAATGGCTGGCATGACATCATG	11885		
QY	541	TTCCAGCTGGAGTCTTTATGCCCCCTGGCATCATCTTATTTTGTCTCTTCAAGATTGTT	600		
DB	11886	TTCCAGCTGGAGTCTTTATGCCCCCTGGCATCATCTTATTTTGTCTCTTCAAGATTGTT	11945		
QY	601	TGGAGCTGAGGGGAGGACAGCTGGCCAGACAGGCTGGATGAAGAGGAGGCCCGG	660		
DB	11946	TGGAGCTGAGGGGAGGACAGCTGGCCAGACAGGCTGGATGAAGAGGAGGCCCGG	12005		
QY	661	TTCAATGTTGGTGGCAATTTGTTTCATCACAATGCTACCTGCCAGCGTGTCTGTAGA	720		
DB	12006	TTCAATGTTGGTGGCAATTTGTTTCATCACAATGCTACCTGCCAGCGTGTCTGTAGA	12065		
QY	721	CTCTATTCTCTGGAGGGTGGCCCTCGAGTGCCTGGATCCCTCTGTCCATGGGCGCTG	780		
DB	12066	CTCTATTCTCTGGAGGGTGGCCCTCGAGTGCCTGGATCCCTCTGTCCATGGGCGCTG	12125		
QY	781	CACATAACCCCTCAGCTTACCTACATGAACAGCATGCTGGATCCCTGGTGTATTATTTT	840		

Db	12126																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
----	-------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

```
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Unpublished
2 (bases 1 to 204062)
Worley, K.C.
Direct Submission
Submitted (22-MAR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Sep 30, 2001 this sequence version replaced gi:14861654.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HAOK
Center clone name: RP11-324E6
----- Summary Statistics
Sequencing vector: Plasmid; M77789
Sequencing vector: M13; L08821
Chemistry: Dye-terminator Big Dye; 92% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 203977 bases at least Q40
Consensus quality: 208562 bases at least Q30
Consensus quality: 210152 bases at least Q20
Estimated insert size: 205209; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-gel estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 50446: contig of 50446 bp in length
* 50447 50546: gap of unknown length
* 50547 93563: contig of 43017 bp in length
* 93564 125647: gap of unknown length
* 125648 125647: contig of 28884 bp in length
* 125648 122647: gap of unknown length
* 122648 140900: contig of 18253 bp in length
* 140901 141000: gap of unknown length
* 141001 152275: contig of 11275 bp in length
* 152276 152375: gap of unknown length
* 152376 163676: contig of 11301 bp in length
* 163677 163776: gap of unknown length
* 163777 173214: contig of 9438 bp in length
* 173215 173314: gap of unknown length
* 173315 181113: contig of 7799 bp in length
* 181114 181213: gap of unknown length
* 181214 181717: contig of 5964 bp in length
* 181718 187277: gap of unknown length
* 187278 192372: contig of 5095 bp in length
* 192373 192473: gap of unknown length
* 192473 196852: contig of 4380 bp in length
* 196853 200822: gap of unknown length
* 200823 200923: contig of 3870 bp in length
* 200923 204062: contig of 3140 bp in length.
Location/Qualifiers
1. 204062
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-324E6"
53813 a 45305 c 44679 g 59056 t 1209 others
BASE COUNT
ORIGIN

Query Match 100.0%; Score 1041; DB 2; Length 204062;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTACAACGGGTGCTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 60
|||||
Db 15571 ATGTACAACGGGTGCTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 15630
|||||
Qy 61 CTGCTCATTTGTCGCTTTGTCGGGCGCACTAGGCAATGGGGTGGCCCTGTGTGGTTTC 120
|||||
Db 15631 CTGCTCATTTGTCGCTTTGTCGGGCGCACTAGGCAATGGGGTGGCCCTGTGTGGTTTC 15690
|||||
Qy 121 TGCCTCCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTTCAATTTGGCCGTGGCT 180
|||||
Db 15691 TGCCTCCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTTCAATTTGGCCGTGGCT 15750
|||||
Qy 181 GATTTCTCTTATGATCTGCCCTGCTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
|||||
Db 15751 GATTTCTCTTATGATCTGCCCTGCTTTCGGACAGACTATTACCTCAGACGTAGACAC 15810
|||||
Qy 241 TGGGCTTTTGGGGACATTCCTCGCCGAGTGGGGCTCTTACGTTGGCCATGAACAGGGCC 300
|||||
Db 15811 TGGGCTTTTGGGGACATTCCTCGCCGAGTGGGGCTCTTACGTTGGCCATGAACAGGGCC 15870
|||||
Qy 301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC 360
|||||
Db 15871 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC 15930
|||||
Qy 361 CACCACGGGTGAACACTATCTCCACCGGGTGGCGCTGGCATCTCTGCACCTGTGG 420
|||||
Db 15931 CACCACGGGTGAACACTATCTCCACCGGGTGGCGCTGGCATCTCTGCACCTGTGG 15990
|||||
Qy 421 GCCCTGTGTCATCTCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGAAGAG 480
|||||
Db 15991 GCCCTGTGTCATCTCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGAAGAG 16050
|||||
Qy 481 ACGCCGCTCTCTGTGAGAGTTCATCATGAGTGGGCAATGGCTGCATGACATCATG 540
|||||
Db 16051 ACGCCGCTCTCTGTGAGAGTTCATCATGAGTGGGCAATGGCTGCATGACATCATG 16110
|||||
Qy 541 TTCCAGCTGGAGTCTTTATGCCCTCGGCATCATCTATTTTTCTCTTCAAGATTCTT 600
|||||
Db 16111 TTCCAGCTGGAGTCTTTATGCCCTCGGCATCATCTATTTTTCTCTTCAAGATTCTT 16170
|||||
Qy 601 TGGAGCTTGAGCGGAGGAGCAGCTGGCCAGACAGAGTTCGGATGAAGAGGCGACCCGG 660
|||||
Db 16171 TGGAGCTTGAGCGGAGGAGCAGCTGGCCAGACAGAGTTCGGATGAAGAGGCGACCCGG 16230
|||||
Qy 661 TTCATCATGTTGGTGGCAATTTGTTTCATCATGCTTACCTGCCCCAGGGTGTCTGTCTAGA 720
|||||
Db 16231 TTCATCATGTTGGTGGCAATTTGTTTCATCATGCTTACCTGCCCCAGGGTGTCTGTCTAGA 16290
|||||
Qy 721 CTCTATTTCCTCTGGAGGGTCCCTCGAGTGGCTGGCATCCCTCTGTCCATGGGGCCCTG 780
|||||
Db 16291 CTCTATTTCCTCTGGAGGGTCCCTCGAGTGGCTGGCATCCCTCTGTCCATGGGGCCCTG 16350
|||||
Qy 781 CACATAACCCCTCAGCTCACCTACATGAACAGCATGCTGGATCCCTCGTGTATATTATTT 840
|||||
Db 16351 CACATAACCCCTCAGCTCACCTACATGAACAGCATGCTGGATCCCTCGTGTATATTATTT 16410
|||||
Qy 841 TCAAGCCCTCTCTTTCCCAAAATTTACAAAGCTCAAAATCTCGAGTCTGAAACCCCAAG 900
|||||
Db 16411 TCAAGCCCTCTCTTTCCCAAAATTTACAAAGCTCAAAATCTCGAGTCTGAAACCCCAAG 16470
|||||
Qy 901 CAGCCAGGACACTCAAAACACAAAGCCGGAAGAGATGCCAATTTGCAACCTCGGTGCG 960
|||||
Db 16471 CAGCCAGGACACTCAAAACACAAAGCCGGAAGAGATGCCAATTTGCAACCTCGGTGCG 16530
|||||
Qy 961 AGGAGTTGCATCAGTGTGGCAAAATAGTTTCCAAAGCCAGTCTGATGGCAATGGATCCC 1020
|||||
Db 16531 AGGAGTTGCATCAGTGTGGCAAAATAGTTTCCAAAGCCAGTCTGATGGCAATGGATCCC 16590
|||||
Qy 1021 CACATTGTTGAGTGGCACTGA 1041
```

|||||
Db 16591 CACATTGTTGAGTGGCACTGA 16611

RESULT 12

AX338374

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 1039; Conservative

0; Mismatches

2; Indels

0; Gaps

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

|||||
Db 16591 CACATTGTTGAGTGGCACTGA 16611

RESULT 12

AX338374

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 1039; Conservative

0; Mismatches

2; Indels

0; Gaps

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

|||||
Db 16591 CACATTGTTGAGTGGCACTGA 16611

RESULT 12

AX338374

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 1039; Conservative

0; Mismatches

2; Indels

0; Gaps

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

|||||
Db 16591 CACATTGTTGAGTGGCACTGA 16611

RESULT 12

AX338374

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 1039; Conservative

0; Mismatches

2; Indels

0; Gaps

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

|||||
Db 16591 CACATTGTTGAGTGGCACTGA 16611

RESULT 12

AX338374

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 1039; Conservative

0; Mismatches

2; Indels

0; Gaps

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

|||||
Db 16591 CACATTGTTGAGTGGCACTGA 16611

RESULT 12

AX338374

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 1039; Conservative

0; Mismatches

2; Indels

0; Gaps

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;


```
/db_xref="taxon:9606"
/cell_line="human monocyte"
/cell_type="monocyte"
61..1224
/gene="HM74"
61..1224
/gene="HM74"
/codon_start=1
/product="HM74"
/protein_id="BAA01721.1"
/db_xref="GI:219867"
/translation="MNRHLDHLEIDKNCVRRDFIAKVLPPVLGFIRGLLG
NGLAWIFCFHLKSKSRIFENLAVADFLIICLPVMDIYVRRDWNFGDIPCLR
VLFFAMNRQSIIFTVVAVRYFRVPHHANKISNMTAAIISCLLWITVGLTV
HLKKKLIQNGPANCISFISCHTFRWEAMFLLEFLPLIGLIFCSARIWISLQR
QMDRAKTIKTFMVAIVFVICFLPSVVVIRIFWLLHTSGTQNCVRSVDLAF
FITLSFTYMSMLDPVVVYFSSPFPNPFSTLINRCLORKMTGPEDNNRSTSVELTGD
PNKTRGAPEALMANGSPSPSYLGPISNNHKKGHCHQEPASLEKQLGCCIE"
BASE COUNT 470 a 530 c 525 g 526 t
ORIGIN

Query Match 3.1%; Score 32; DB 9; Length 2051;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 790 CTCAGCTTCACCTACATGAACAGCATGCTGGA 821
|||||
Db 898 CTCAGCTTCACCTACATGAACAGCATGCTGGA 929
|||||

RESULT 17
AXI47888
LOCUS AXI47888 24 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 133 from Patent WO0136473.
ACCESSION AXI47888
VERSION AXI47888.1 GI:14346883
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequence.
REFERENCE 1 (bases 1 to 24)
AUTHORS Vogeli,G., Wood,L.S., Parodi,L.A., Hiebsch,R.R., Lind,P.,
Slightom,J., Schellin,K.A., Kaytes,P.S., Bannigan,C.M., Ruff,V.,
Sejltz,T. and Huff,R.M.
TITLE Novel g protein-coupled receptors
JOURNAL Patent: WO 0136473-A 133 25-MAY-2001;
PHARMACIA & UPJOHN COMPANY (US)
FEATURES
source
1..24
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Novel Sequence"
BASE COUNT 3 a 7 c 5 g 9 t
ORIGIN

Query Match 2.3%; Score 24; DB 6; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 707 GCGTGTCTGCTAGACTCTATTTC 730
|||||
Db 1 GCGTGTCTGCTAGACTCTATTTC 24
|||||

RESULT 18
AXI47917
LOCUS AXI47917 24 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 162 from Patent WO0136473.
ACCESSION AXI47917
VERSION AXI47917.1 GI:14346912
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequence.
REFERENCE 1 (bases 1 to 52)
AUTHORS Vogeli,G., Wood,L.S., Parodi,L.A., Hiebsch,R.R., Lind,P.,
Slightom,J., Schellin,K.A., Kaytes,P.S., Bannigan,C.M., Ruff,V.,
Sejltz,T. and Huff,R.M.
TITLE Novel g protein-coupled receptors
JOURNAL Patent: WO 0136473-A 133 25-MAY-2001;
PHARMACIA & UPJOHN COMPANY (US)
FEATURES
source
1..24
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Novel Sequence"
BASE COUNT 3 a 7 c 5 g 9 t
ORIGIN

Query Match 2.3%; Score 24; DB 6; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 707 GCGTGTCTGCTAGACTCTATTTC 730
|||||
Db 1 GCGTGTCTGCTAGACTCTATTTC 24
|||||

RESULT 19
AXI47885
LOCUS AXI47885 52 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 130 from Patent WO0136473.
ACCESSION AXI47885
VERSION AXI47885.1 GI:14346880
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequence.
REFERENCE 1 (bases 1 to 52)
AUTHORS Vogeli,G., Wood,L.S., Parodi,L.A., Hiebsch,R.R., Lind,P.,
Slightom,J., Schellin,K.A., Kaytes,P.S., Bannigan,C.M., Ruff,V.,
Sejltz,T. and Huff,R.M.
TITLE Novel g protein-coupled receptors
JOURNAL Patent: WO 0136473-A 130 25-MAY-2001;
PHARMACIA & UPJOHN COMPANY (US)
FEATURES
source
1..52
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Novel Sequence"
BASE COUNT 12 a 14 c 12 g 14 t
ORIGIN

Query Match 2.3%; Score 24; DB 6; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 707 GCGTGTCTGCTAGACTCTATTTC 730
|||||
Db 1 GCGTGTCTGCTAGACTCTATTTC 24
|||||

RESULT 20
AXI47914
LOCUS AXI47914 52 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 159 from Patent WO0136473.
ACCESSION AXI47914
VERSION AXI47914.1 GI:14346909
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequence.
REFERENCE 1 (bases 1 to 52)
AUTHORS Vogeli,G., Wood,L.S., Parodi,L.A., Hiebsch,R.R., Lind,P.,
Slightom,J., Schellin,K.A., Kaytes,P.S., Bannigan,C.M., Ruff,V.,
Sejltz,T. and Huff,R.M.
```


/note="129/SVJ ES cell, BAC clone"

2404..2461

/gene="Puma-g"

2404..4323

/gene="Puma-g"

2462..3544

/gene="Puma-g"

/codon_start=1

/product="putative seven transmembrane spanning receptor"

/protein_id="CAC17791.1"

/db_xref="GI:11558406"

/translation="MSKSDHFLVINGKCCVFRDENIAKVLPPVLGLFEVFLGLNGL
ALWIFCFLKSKSRIFLFLAVADFLIIICLPFLTDNYVHNWDRGGIPCRVWLF
MLAMRQSGIIFLTVAVDYRVVPHHFLNKNRTAAIISCFLWGLTIGLTVHLL
YTNMTKNEAIVCSSEICYNFRWHDAMFLLEFLPLAIILFCSGRIIWSLRQROMD
RHAKIKRAINFYVAIVFIICFLPSVAVRIRIFWLLKYVNRNDIYSSVDLAFTT
LSFTFGALMDPVYIFSPFPNFSTCINRCLRKLTGEPDNNRSTSVELTGDPST
3545..4323

/gene="Puma-g"

4211..4216

/gene="Puma-g"

BASE COUNT 1921 a 1619 c 1670 g 1890 t

ORIGIN

Query Match 2.2%; Score 23; DB 10; Length 7100;

Best Local Similarity 100.0%; Pred. No. 0.84; Mismatches 0; Indels 0; Gaps 0;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 799 ACCTACATGAACAGCATGCTGGA 821

|||||

Db 3299 ACCTACATGAACAGCATGCTGGA 3321

RESULT 24

AL390918

LOCUS Human DNA sequence from clone Rp11-569E4 on chromosome 13, complete

DEFINITION

sequence.

ACCESSION AL390918

VERSION AL390918.7

KEYWORDS GI:13157583

SOURCE HTG.

ORGANISM human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 83550)

Mashreghi-Mohammadi, M.

Direct Submission

Submitted (22-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

requests: clonerequest@sanger.ac.uk

On Feb 27, 2001 this sequence version replaced gi:13121436.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest. The following

abbreviations are used to associate primary accession numbers given

in the feature table with their source databases: Em: EMBL; Sw:

SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP

database can be found at

http://www.sanger.ac.uk/Projects/C.elegans/wormpep This sequence

was generated from part of bacterial clone contigs of human

chromosome 13, constructed by the Sanger Centre Chromosome 13

Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr13

Rp11-569E4 is from the library RPCI-11.2 constructed by the group

of Pieter de Jong. For further details see

http://www.chori.org/bacpac/home.htm

VECTOR: pBACE3.6

IMPORTANT: This sequence is not the entire insert of clone

Rp11-569E4. It may be shorter because we sequence overlapping

sections only once, except for a 100 base overlap.

The true left end of clone Rp11-322A14 is at 83451 in this

sequence. The true right end of clone Rp11-141M24 is at 100 in this

sequence.

FEATURES

source

1..83550

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="13"

/clone="Rp11-569E4"

/clone_11b="RPCI-11.2"

/complement(1..88)

/note="match: GSS: Em:AQ387473"

/complement(1..100)

/note="match: GSS: Em:AQ052321 Em:AQ386521"

107..620

/note="match: GSS: Em:AQ172927"

1341..1626

/note="AluX repeat: matches 1..292 of consensus"

2400..2598

/note="MER6 repeat: matches -65..864 of consensus"

3013..3385

/note="THE1B repeat: matches 1..364 of consensus"

4683..4904

/note="MIR repeat: matches 1..238 of consensus"

5283..5363

/note="9 copies 9 mer aatatttt 67% conserved"

5771..6098

/note="AluX repeat: matches 1..299 of consensus"

complement(6378..6805)

/note="match: GSS: Em:AQ231235"

complement(6397..6794)

/note="match: GSS: Em:AQ011789"

6850..7320

/note="match: GSS: Em:AQ560705"

complement(7341..7762)

/note="match: GSS: Em:A2516719"

7425..7474

/note="25 copies 2 mer tg 94% conserved"

7428..7475

/note="12 copies 4 mer tgt 95% conserved"

8228..8535

/note="AluSp repeat: matches 1..309 of consensus"

8901..9049

/note="MER5A repeat: matches 9..184 of consensus"

complement(9463..9928)

/note="match: GSS: Em:AQ388399"

9500..10559

/note="MER11C repeat: matches 1..1057 of consensus"

11548..11884

/note="MLT1C repeat: matches 108..466 of consensus"

11891..12532

/note="MER44 repeat: matches 1..728 of consensus"

complement(14281..14711)

/note="match: GSS: Em:AQ252887"

complement(14285..14711)

/note="match: GSS: Em:AQ179843"

complement(14350..14638)

/note="match: GSS: Em:B89417"

complement(14414..14678)

/note="match: GSS: Em:AQ618066"

complement(14461..14614)

/note="match: GSS: Em:B86556"

14828..14883

/note="2 copies 28 mer 96% conserved"

```

repeat_region 15047..15207
/notes="MIR repeat: matches 36..198 of consensus"
repeat_region 17778..18088
/notes="AluX repeat: matches 1..312 of consensus"
misc_feature 18567..19051
/notes="match: GSS: Em:AQ322687"
repeat_region 18874..19000
/notes="L1M4 repeat: matches 4074..4207 of consensus"
repeat_region 20121..20152
/notes="match: GSS: Em:AQ322687"
repeat_region 20391..20695
/notes="L2 repeat: matches 2249..2578 of consensus"
repeat_region 20850..21167
/notes="L2 repeat: matches 1701..2046 of consensus"
misc_feature 21649..22168
/notes="match: GSS: Em:AQ589986"
repeat_region 21897..21946
/notes="25 copies 2 mer gt 78% conserved"
repeat_region 21902..21945
/notes="11 copies 4 mer tgtg 79% conserved"
repeat_region 22767..23130
/notes="MER57A repeat: matches 1..369 of consensus"
repeat_region 23134..23340
/notes="23 copies 9 mer atattattt 58% conserved"
repeat_region 23743..24028
/notes="AluX repeat: matches 1..295 of consensus"
misc_feature complement(23915..24399)
/notes="match: GSS: Em:AQ595323"
misc_feature complement(23993..24393)
/notes="match: GSS: Em:AQ201248"
misc_feature complement(24032..24391)
/notes="match: GSS: Em:AQ538244"
repeat_region 24404..25001
/notes="match: GSS: Em:AQ346571"
repeat_region 24434..24511
/notes="39 copies 2 mer aa 70% conserved"
repeat_region 25296..25373
/notes="MIR repeat: matches 177..258 of consensus"
repeat_region 26106..26147
/notes="2 copies 21 mer 100% conserved"
repeat_region 27463..27623
/notes="MIR repeat: matches 61..229 of consensus"
repeat_region 27743..28057
/notes="AluX repeat: matches 1..312 of consensus"
repeat_region 28846..29170
/notes="MER46C repeat: matches 4..335 of consensus"
repeat_region 29342..29509
/notes="6 copies 28 mer 70% conserved"
repeat_region 29350..29517
/notes="8 copies 21 mer 66% conserved"
repeat_region 29363..29522
/notes="80 copies 2 mer tc 68% conserved"
repeat_region 29563..29844
/notes="AluX repeat: matches 1..282 of consensus"
repeat_region 30820..30988
/notes="MIR repeat: matches 17..193 of consensus"
repeat_region 31927..32118
/notes="MIR repeat: matches 57..261 of consensus"
repeat_region 34618..35081
/notes="L2 repeat: matches 1796..2267 of consensus"
repeat_region 35152..35237
/notes="L2 repeat: matches 2403..2491 of consensus"
repeat_region 36203..36224
/notes="11 copies 2 mer ag 100% conserved"
misc_feature 36855..37383
/notes="match: GSS: Em:AQ390049"
misc_feature complement(37055..37577)
/notes="match: GSS: Em:AQ670939"
repeat_region 38649..38857
/notes="MER20 repeat: matches 21..218 of consensus"
misc_feature complement(39901..40653)
/notes="match: GSS: Em:AQ541398"
misc_feature complement(40197..40653)
/notes="match: GSS: Em:AQ349212"
40677..40951
/notes="match: GSS: Em:AQ543863"
complement(41575..41939)
/notes="match: GSS: Em:AQ100514"
41931..42003
/notes="L2 repeat: matches 2488..2570 of consensus"
42045..42879
/notes="match: GSS: Em:AQ748881"
42046..42524
/notes="match: GSS: Em:B82390"
42575..42618
/notes="11 copies 4 mer gtgt 84% conserved"
42686..45177
/notes="L1PA16 repeat: matches 3689..6157 of consensus"
45178..45245
/notes="17 copies 4 mer cctt 98% conserved"

Query Match 2.1% Score 22; DB 9; Length 83550;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 860 AATTCTACACAGCTCAAAAT 881
|||||
DB 1317 AATTCTACACAGCTCAAAAT 1338

RESULT 25
AC012158/c
LOCUS AC012158.30 GI:17402749
DEFINITION Homo sapiens chromosome 12 clone RP11-290121, WORKING DRAFT
ACCESSION AC012158
VERSION AC012158.30
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 152664)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alzbrooks,S.L., Amarante,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Bivona,M., Brown,E., Brown,M., Bryant,N.P., Buha,C.,
Bowie,S., Brieva,M., Burrell,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulsegh,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S.,
Ogulu,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoshitari,N.,
Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,

```

Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 152664)
Worley, K.C.

Direct Submission
Submitted (21-Oct-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Dec 7, 2001 this sequence version replaced gi:17155006.

COMMENT

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HMOA
Center clone name: RP11-290I21
----- Summary Statistics
Sequencing vector: Plasmid; M7789
Chemistry: Dye-primer Bodipy: 2% of reads
Chemistry: Dye-terminator Big Dye: 98% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 152317 bases at least Q40
Consensus quality: 154222 bases at least Q30
Consensus quality: 155780 bases at least Q20
Estimated insert size: 155036; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 8.8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 83772: contig of 83772 bp in length
* 83773 83872: gap of unknown length
* 83873 152664: contig of 68792 bp in length.

FEATURES

source

1. .152664
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-290I21"

BASE COUNT 44420 a 31059 c 31169 g 45911 t 105 others
ORIGIN

Query Match 2.1%; Score 22; DB 2; Length 152664;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 901 CAGCCAGGACATCAAAACAC 922

|||||
Db 144495 CAGCCAGGACATCAAAACAC 144474

RESULT 26

AC011080

LOCUS

AC011080 193735 bp DNA linear HTG 26-OCT-2000
DEFINITION Homo sapiens clone RP11-45K2, WORKING DRAFT SEQUENCE, 2 ordered
pieces.

ACCESSION

AC011080

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

AC011080.3 GI:11024919

HTG; HTGS_PHASE2; HTGS_DRAFT.

human.

Homo sapiens

Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 183735)

Birren, B., Linton, L., Nusbaum, C. and Lander, E.

Homo sapiens, clone RP11-45K2

Unpublished

2 (bases 1 to 183735)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Balwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Bouckhalter, B., Brown, A., Castelle, A., Collings, S., Collins, S., Conliffe, A., Cooke, P., DeArelano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lehoczy, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Melidri, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

Direct Submission

TITLE

JOURNAL

COMMENT

Submitted (01-Oct-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Oct 26, 2000 this sequence version replaced gi:7341694.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L1174

Center clone name: 45_K_2

----- Summary Statistics

Sequencing vector: M13; M77815; 44% of reads

Sequencing vector: Plasmid; n/a; 56% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 182296 bases at least Q40

Consensus quality: 182874 bases at least Q30

Consensus quality: 183232 bases at least Q20

Insert size: 188000; agarose-fp

Insert size: 183635; sum-of-contigs

Quality coverage: 7.7 in Q20 bases; agarose-fp

Quality coverage: 7.9 in Q20 bases.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 2 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 21343: contig of 21343 bp in length

* 21344 21443: gap of 100 bp

* 21444 183735: contig of 162292 bp in length.

FEATURES

source

1. .183735

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="RP11-45K2"

/clone_lib="RP11-11 Human Male BAC"

1. .21343

/note="assembly_fragment"

clone_end:SP6

misc_feature

misc_feature vector_side:left
21444..183735
/note=assembly_fragment
clone_end:T7
vector_side:right

BASE COUNT 56460 a 33977 c 35374 g 57820 t 104 others
ORIGIN

Query Match 2.1% Score 22; DB 2; Length 183735;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 860 AATTCTACAACAGCTCAAAAT 881
|||||

Db 160651 AATTCTACAACAGCTCAAAAT 160672

RESULT 27
AC055724/c
LOCUS AC055724 199607 bp DNA linear HTG 31-JAN-2002
DEFINITION Homo sapiens chromosome 12 clone RP11-985023, *** SEQUENCING IN
PROGRESS ***, 27 unordered pieces.
AC055724
VERSION AC055724.10 GI:18449533
KEYWORDS HTG; HTGS_PHASE1.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 199607)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunga,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbata,J., Benton,J., Blmage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brileva,M., Brown,M., Brown,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Devila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denna,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Haves,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
Homs,I., Howard,S., Huber,J., Hulya,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I.,
Sodergren,E., Sonalike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Uman,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wall,R.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williams,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 199607)

AUTHORS
TITLE
JOURNAL
COMMENT

Worley, K.C.
Direct Submission
Submitted (18-APR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jan 31, 2002 this sequence version replaced gi:10086369.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: HAUT
Center clone name: RP11-985023

----- Summary Statistics
Sequencing vector: M13; L08821

Chemistry: Dye-terminator Big Dye: 50% of reads
Chemistry: Dye-terminator Big Dye: 50% of reads

Assembly program: Phrap; version 0.990329First call to
findPhrapList

Consensus quality: 173746 bases at least Q40
Consensus quality: 186262 bases at least Q30
Consensus quality: 192992 bases at least Q20
Estimated insert size: 196552; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-gel estimation
Quality coverage: 2.9x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 27 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 18044: contig of 18044 bp in length
18144: gap of unknown length
18145: contig of 17902 bp in length
36047: gap of unknown length
36146: gap of unknown length
36147: contig of 12578 bp in length
48724: gap of unknown length
48824: gap of unknown length
58725: contig of 9901 bp in length
58825: gap of unknown length
58826: contig of 11190 bp in length
58826: gap of unknown length
70015: gap of unknown length
70016: gap of unknown length
70116: contig of 13699 bp in length
83814: contig of 13699 bp in length
83914: gap of unknown length
95389: contig of 11475 bp in length
95489: gap of unknown length
107170: contig of 11681 bp in length
107170: gap of unknown length
107270: contig of 8918 bp in length
116188: gap of unknown length
116189: gap of unknown length
116288: contig of 8948 bp in length
125236: gap of unknown length
125237: gap of unknown length
125337: contig of 9004 bp in length
134340: gap of unknown length
134341: gap of unknown length
139657: contig of 5217 bp in length
139657: gap of unknown length
139658: contig of 6180 bp in length
145937: gap of unknown length
145938: gap of unknown length
151114: contig of 5077 bp in length
151115: gap of unknown length
151215: contig of 4370 bp in length
155585: gap of unknown length
155585: contig of 6326 bp in length
162010: gap of unknown length
162011: contig of 4911 bp in length
167021: gap of unknown length
167022: gap of unknown length
167121: contig of 4558 bp in length
171679: contig of 4558 bp in length
171680: gap of unknown length

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE

1 (sites)
Sasaki,T., Matsumoto,T. and Yamamoto,K.
Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC
clone:P0622F03

JOURNAL

Published Only in DataBase (2001) In press

REFERENCE

2 (bases 1 to 175644)
Sasaki,T., Matsumoto,T. and Yamamoto,K.

JOURNAL

Submitted (20-JUN-2001) Takuji Sasaki, National Institute of
Agrobiological Resources, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@agr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)

COMMENT

NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES

source

1..175644
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosomes="6"
/clone="P0622F03"

BASE COUNT 46662 a 41378 c 40541 g 46713 t 350 others
ORIGIN

Query Match 2.0%; Score 21; DB 2: Length 175644;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 560 GTTCATCATGGTGGTGGCAAT 580

|||||
Db 51809 GTTCATCATGGTGGTGGCAAT 51789

RESULT 33

AC006499

LOCUS AC006499 195384 bp DNA linear PRI 02-NOV-1999
DEFINITION Homo sapiens chromosome 4 clone C0494H11, complete sequence.

AC006499

AC006499.13 GI:6175133

KEYWORDS

HTG.

SOURCE

human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 195384)
Stone,N.E., Schmutz,J.J., Shang,J., Pennacchio,L.A., Cox,D.R. and
Myers,R.M.

TITLE

Direct Submission

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 195384)
Stone,N.E., Schmutz,J.J., Cox,D.R. and Myers,R.M.

TITLE

Direct Submission

JOURNAL

Submitted (03-FEB-1999) Department of Genetics, Stanford Human
Genome Center, 855 California Avenue, Palo Alto, CA 94304, USA

REFERENCE

3 (bases 1 to 195384)
Stone,N.E., Schmutz,J.J., Shang,J., Pennacchio,L.A., Cox,D.R. and
Myers,R.M.

TITLE

Direct Submission

JOURNAL

Submitted (02-NOV-1999) Department of Genetics, Stanford Human
Genome Center, 4005 Miranda Avenue, Palo Alto, CA 94304, USA

COMMENT

On Nov 2, 1999 this sequence version replaced gi:5757494.

Quality: Phrap Quality >=40 99.7% of Sequence;
Estimated Total Number of Errors is 0.6.

STS Content:

SHGC-86135 G61558
WI-9872 G05448
SHGC-50686 G33724
SHGC-24200 G33354
SHGC-51172 G34253
WI-4297 G04613
SHGC-51247 G34292
SHGC4-1530 G01882
WI-945 G02751
SHGC-50683 G33721
SHGC-84465 G51855
SHGC-82103 G51922

FEATURES

source

1..195384
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosomes="4"
/clone="C0494H11"
/clone_lib="ROSWELL PARK CANCER RPCI - 11 Human Male BAC
Library"

BASE COUNT 58996 a 44583 c 41620 g 50185 t
ORIGIN

Query Match 2.0%; Score 21; DB 9: Length 195384;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 CACCATCTCCAGGTGATGCC 56

|||||
Db 68375 CACCATCTCCAGGTGATGCC 68395

RESULT 34

AC024933/c

LOCUS

AC024933 317285 bp DNA linear HTG 06-FEB-2002
DEFINITION Homo sapiens chromosome 3 clone RP11-219D15, WORKING DRAFT
SEQUENCE, 27 unordered pieces.

AC024933

AC024933.28 GI:18542847

KEYWORDS

HTG.

SOURCE

human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 317285)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louiseged,H.,
Lozadó,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,

Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Fu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E., Sotniko,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Umanik,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 317285)
Worley,K.C.

Direct Submission
Submitted (03-MAR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Feb 6, 2002 this sequence version replaced gi:15723405.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HAIX
Center clone name: RP11-219D15
----- Summary Statistics
Sequencing vector: Plasmid;
Sequencing vector: MJ3;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 299405 bases at least Q40
Consensus quality: 309299 bases at least Q30
Consensus quality: 315960 bases at least Q20
Estimated insert size: 325233; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-gel estimation
Quality coverage: 7.9x in Q20 bases; sum-of-contigs estimation

COMMENT

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 27 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2040: contig of 2040 bp in length
* 2041 2140: gap of unknown length
* 2141 4175: contig of 2035 bp in length
* 4176 4275: gap of unknown length
* 4276 6607: contig of 2332 bp in length
* 6608 6708: gap of unknown length
* 6708 8724: contig of 2017 bp in length
* 8725 8825: gap of unknown length
* 8825 12042: contig of 3218 bp in length
* 12043 12142: gap of unknown length
* 12143 14350: contig of 2208 bp in length
* 14351 14451: gap of unknown length
* 14451 17189: contig of 2739 bp in length
* 17190 17289: gap of unknown length
* 17290 20516: contig of 3227 bp in length
* 20517 20616: gap of unknown length
* 20616 23616: contig of 3000 bp in length
* 23617 23716: gap of unknown length

* 23717 27685: contig of 3969 bp in length
* 27686 27785: gap of unknown length
* 27786 31495: contig of 3710 bp in length
* 31496 31595: gap of unknown length
* 31596 36288: contig of 4693 bp in length
* 36289 36389: gap of unknown length
* 36389 39485: contig of 3097 bp in length
* 39486 39585: gap of unknown length
* 39586 44802: contig of 5217 bp in length
* 44803 44903: gap of unknown length
* 44903 49767: contig of 4865 bp in length
* 49768 49868: gap of unknown length
* 49868 55018: contig of 5151 bp in length
* 55019 55118: gap of unknown length
* 55119 60785: contig of 5668 bp in length
* 60786 60885: gap of unknown length
* 60886 67684: contig of 6798 bp in length
* 67685 67785: gap of unknown length
* 67785 74694: contig of 6910 bp in length
* 74695 74794: gap of unknown length
* 74795 82085: contig of 7292 bp in length
* 82087 82187: gap of unknown length
* 82188 88698: contig of 6512 bp in length
* 88699 88798: gap of unknown length
* 88799 97204: contig of 8406 bp in length
* 97205 97304: gap of unknown length
* 97305 105093: contig of 7789 bp in length
* 105094 105193: gap of unknown length
* 105194 115204: contig of 10011 bp in length
* 115205 115304: gap of unknown length
* 115305 126830: contig of 11526 bp in length
* 126831 126930: gap of unknown length
* 126931 147531: contig of 20601 bp in length
* 147532 147631: gap of unknown length
* 147632 317285: contig of 169654 bp in length.

FEATURES

source

1. 317285
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-219D15"
BASE COUNT 83820 a 72967 c 70387 g 87453 t 2658 others
ORIGIN

Query Match 2.0% Score 21; DB 2; Length 317285;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 614 GGAGGCAGCAGCTGCCAGAC 634
|||||
Db 247193 GGAGGCAGCAGCTGCCAGAC 247173

RESULT 35

AX147864
LOCUS AX147864 20 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 109 from Patent W00136473.
ACCESSION AX147864
VERSION AX147864.1 GI:14346860
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Vogel,G., Wood,L.S., Parodi,L.A., Hiebsch,R.R., Lind,P., Slightom,J., Schellin,K.A., Kaytes,P.S., Bannigan,C.M., Ruff,V., Sejlitz,T. and Huff,R.M.
TITLE Novel g protein-coupled receptors
JOURNAL Patent: WO 0136473-A 109 25-MAY-2001;
PHARMACIA & UPJOHN COMPANY (US)
FEATURES Location/Qualifiers
source 1. .20
/organism="synthetic construct"

BASE COUNT 5 a 7 c 4 g 4 t
ORIGIN

Query Match 1.9%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 GAAGCCCGACGACGTGTACC 160
|||||
Db 1 GAAGCCCGACGACGTGTACC 20

RESULT 36
AX147865/c
LOCUS AX147865 20 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 110 from Patent WO0136473.
ACCESSION AX147865
VERSION AX147865.1 GI:14346861
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 20)
AUTHORS Vogeli,G., Wood,L.S., Parodi,L.A., Hiebsch,R.R., Lind,P.,
Slightom,J., Schellin,K.A., Kaytes,P.S., Bannigan,C.M., Ruff,V.,
Sejltiz,T. and Huff,R.M.
TITLE Novel g protein-coupled receptors
JOURNAL Patent: WO 0136473-A 110 25-MAY-2001;
PHARMACIA & UPJOHN COMPANY (US)
FEATURES Location/Qualifiers
source 1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Novel Sequence"

BASE COUNT 5 a 7 c 4 g 4 t
ORIGIN

Query Match 1.9%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 327 GGCTGCGGACAGGTATTCA 346
|||||
Db 20 GGCTGCGGACAGGTATTCA 1

RESULT 37
AX147886/c
LOCUS AX147886 20 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 131 from Patent WO0136473.
ACCESSION AX147886
VERSION AX147886.1 GI:14346881
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 20)
AUTHORS Vogeli,G., Wood,L.S., Parodi,L.A., Hiebsch,R.R., Lind,P.,
Slightom,J., Schellin,K.A., Kaytes,P.S., Bannigan,C.M., Ruff,V.,
Sejltiz,T. and Huff,R.M.
TITLE Novel g protein-coupled receptors
JOURNAL Patent: WO 0136473-A 131 25-MAY-2001;
PHARMACIA & UPJOHN COMPANY (US)
FEATURES Location/Qualifiers
source 1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Novel Sequence"

BASE COUNT 5 a 8 c 3 g 4 t
ORIGIN

Query Match 1.9%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 962 GGAGTTGCATCAGTGGCA 981
|||||
Db 20 GGAGTTGCATCAGTGGCA 1

RESULT 38
AX147915/c
LOCUS AX147915 20 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 160 from Patent WO0136473.
ACCESSION AX147915
VERSION AX147915.1 GI:14346910
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 20)
AUTHORS Vogeli,G., Wood,L.S., Parodi,L.A., Hiebsch,R.R., Lind,P.,
Slightom,J., Schellin,K.A., Kaytes,P.S., Bannigan,C.M., Ruff,V.,
Sejltiz,T. and Huff,R.M.
TITLE Novel g protein-coupled receptors
JOURNAL Patent: WO 0136473-A 160 25-MAY-2001;
PHARMACIA & UPJOHN COMPANY (US)
FEATURES Location/Qualifiers
source 1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Novel Sequence"

BASE COUNT 5 a 8 c 3 g 4 t
ORIGIN

Query Match 1.9%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 962 GGAGTTGCATCAGTGGCA 981
|||||
Db 20 GGAGTTGCATCAGTGGCA 1

RESULT 39
AX147887/c
LOCUS AX147887 48 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 132 from Patent WO0136473.
ACCESSION AX147887
VERSION AX147887.1 GI:14346882
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 48)
AUTHORS Vogeli,G., Wood,L.S., Parodi,L.A., Hiebsch,R.R., Lind,P.,
Slightom,J., Schellin,K.A., Kaytes,P.S., Bannigan,C.M., Ruff,V.,
Sejltiz,T. and Huff,R.M.
TITLE Novel g protein-coupled receptors
JOURNAL Patent: WO 0136473-A 132 25-MAY-2001;
PHARMACIA & UPJOHN COMPANY (US)
FEATURES Location/Qualifiers
source 1..48
/organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT 14 a 15 c 10 g 9 t
ORIGIN

Query Match 1.9%; Score 20; DB 6; Length 48;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 962 GGAGTTGCATCAGTGGCA 981
|||||

```
Db 48 GGAGTTGCATCAGTGTGGCA 29

RESULT 40
AX147916/c
LOCUS AX147916 48 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 161 from Patent WO0136473.
ACCESSION AX147916
VERSION AX147916.1 GI:14346911
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 48)
AUTHORS Vogeli,G., Wood,L.S., Parodi,L.A., Hiebsch,R.R., Lind,P.,
Slightom,J., Schellin,K.A., Kaytes,P.S., Bannigan,C.M., Ruff,V.,
Sejltz,T. and Huff,R.M.
TITLE Novel g protein-coupled receptors
JOURNAL Patent: WO 0136473-A 161 25-MAY-2001;
PHARMACIA & UPJOHN COMPANY (US)
FEATURES
source 1..48
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Novel sequence"
BASE COUNT 14 a 15 c 10 g 9 t
ORIGIN
1
|||||
14 a 15 c 10 g 9 t

Query Match 1.9% Score 20; DB 6; Length 48;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 962 GGAGTTGCATCAGTGTGGCA 981
|||||
Db 48 GGAGTTGCATCAGTGTGGCA 29

RESULT 41
AB023806/c
LOCUS AB023806 2116 bp mRNA linear VRT 25-MAR-2000
DEFINITION Gallus gallus Fz-10 mRNA for Frizzled-10, complete cds.
ACCESSION AB023806
VERSION AB023806.1 GI:5821258
KEYWORDS Frizzled-10.
SOURCE Gallus gallus Stage 18-24 limb bud cDNA to mRNA.
ORGANISM Gallus gallus
REFERENCE 1 (sites)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Kavakami,Y., Wada,N., Nishimatsu,S., Komaguchi,C., Noji,S. and
Nohno,T.
TITLE Identification of chick frizzled-10 expressed in the developing
JOURNAL limb and the central nervous system
MEDLINE Mech. Dev. 91 (1-2), 375-378 (2000)
REFERENCE 20171066
AUTHORS Nohno,T.
TITLE Direct Submission
JOURNAL Submitted (17-FEB-1999) Tsutomu Nohno, Kawasaki Medical School,
Molecular Biology; 577 Matsushima, Kurashiki, Okayama 701-0192,
Japan (E-mail:nohno@bcc.kawasaki-m.ac.jp,
Tel:81-86-462-1111(ex.3501), Fax:81-86-462-1199)
FEATURES
source 1..2116
/organism="Gallus gallus"
/db_xref="taxon:9031"
/tissue_type="limb bud"
/dev_stage="Stage 18-24"
gene 256..2013
CDS 256..2013
/gene="Fz-10"

/note="Involved in Wnt signaling, encoding a Wnt-7a
receptor"
/codon_start=1
/product="Frizzled-10"
/protein_id="BAA83742.1"
/db_xref="GI:5821259"
/translaton="MGPAAGNLVRAVLALCWLAEHCAGISSIDIERPGDRCOPIETP
MCKDIGYNTMRPNLGMHGNQREAAIQLEHFAPLVEYCGHGLKFFLCISLYAPMCTEQ
VSTPIACRVMEQARLKCSPIMEQFNKWPDSLDCKLPKNPNVLCMEAPNNGSD
EPPRGSSMLPMPERPQPSGHDLOQHKDSLRTSCENPKGFHHVEKSASCAPLCTPG
VDVWSKDDKQFAVINTAIWSILICFFSSAFTVLTFLIDPQRFKYPPIFLSNVCYV
YSVGYIIRLFSGAESIACDRDSGLYVIOQEGLESTGCTIVFLVYFPGMASSLWVIL
TLTWFLAAGKKWGHEAIEANSYFHLAAWAIPAVKTIIMLVMRVAGDELGLCYVGS
MDVNALTYGFLVLACVLIIGTSTFSLSGFVALFHRRVMTGGTNTDKLEKLMVRIGV
FSVLYTVPATCVIACYFYERLNDYKIVASQOKCKMNOTKNDLCHMMNNSIAPVEIF
MKVIFMLLVVGIITSGMWITWTKTQSHQNVCSRRLKRSRRKPASVITSSGIYKKPOH
PQXHLAKYESTLQPPYCV"
BASE COUNT 492 a 528 c 544 g 552 t
ORIGIN
1
|||||
492 a 528 c 544 g 552 t

Query Match 1.9% Score 20; DB 5; Length 2116;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 810 CAGCATGCTGGATCCCTGG 829
|||||
Db 762 CAGCATGCTGGATCCCTGG 743

RESULT 42
AF224320/c
LOCUS AF224320 2244 bp mRNA linear VRT 29-MAR-2000
DEFINITION Gallus gallus Frizzled-10 (cfz-10) mRNA, complete cds.
ACCESSION AF224320
VERSION AF224320.1 GI:7340131
KEYWORDS chicken.
SOURCE Gallus gallus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 2244)
AUTHORS Stark,M.R., Biggs,J.J., Schoenwolf,G.C. and Rao,M.S.
TITLE Characterization of Avian Frizzled Genes in Cranial Placode
JOURNAL Development
REFERENCE 2 (bases 1 to 2244)
AUTHORS Stark,M.R., Biggs,J.J., Schoenwolf,G.C. and Rao,M.S.
TITLE Direct Submission
JOURNAL Submitted (13-JAN-2000) Neurobiology and Anatomy, University of
Utah, 50 N. Medical Dr., Salt Lake City, UT 84132, USA
FEATURES
source 1..2244
/organism="Gallus gallus"
/db_xref="taxon:9031"
/dev_stage="stage 4-11 chick"
gene 1..2244
CDS 428..2185
/gene="cfz-10"
/note="7-pass transmembrane receptor"
/codon_start=1
/product="Frizzled-10"
/protein_id="AAF61100.1"
/db_xref="GI:7340132"
/translaton="MGPAAGNLVRAVLALCWLAEHCAGISSIDIERPGDRCOPIETP
MCKDIGYNTMRPNLGMHGNQREAAIQLEHFAPLVEYCGHGLKFFLCISLYAPMCTEQ
VSTPIACRVMEQARLKCSPIMEQFNKWPDSLDCKLPKNPNVLCMEAPNNGSD
EPPRGSSMLPMPERPQPSGHDLOQHKDSLRTSCENPKGFHHVEKSASCAPLCTPG
VDVWSKDDKQFAVINTAIWSILICFFSSAFTVLTFLIDPQRFKYPPIFLSNVCYV
YSVGYIIRLFSGAESIACDRDSGLYVIOQEGLESTGCTIVFLVYFPGMASSLWVIL
TLTWFLAAGKKWGHEAIEANSYFHLAAWAIPAVKTIIMLVMRVAGDELGLCYVGS
MDVNALTYGFLVLACVLIIGTSTFSLSGFVALFHRRVMTGGTNTDKLEKLMVRIGV
```

FSVLTVPATCVIACYFYERLNDYKIVASQOKKMNQTKNLDMMNNSIPAVEIF
MYKIMPLVVGITSGMIWTISKTQLSQWQVCSRRLLKRSRRKPSVITSSGIYKRPQH
POKTHLAKYESTLOPTCV"
BASE COUNT 532 a 573 c 606 g 533 t
ORIGIN

Query Match 1.9%; Score 20; DB 5; Length 2244;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 810 CAGCATGCTGGATCCCTGG 829
|||||
Db 934 CAGCATGCTGGATCCCTGG 915

RESULT 43
HSGROUCH2 12524 bp DNA linear PRI 05-JAN-1999
LOCUS Human grouch protein homolog (AES) gene, exons 2-7 and complete
DEFINITION cds.

ACCESSION U88832
VERSION 1 GI:4099599

KEYWORDS
SEGMENT
SOURCE 2 of 2
human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 12524)
AUTHORS Miyasaka, H., Choudhury, B.K., Hou, E.W. and Li, S.S.

TITLE Molecular cloning and expression of mouse and human CDNA encoding
AES and ESG proteins with strong similarity to Drosophila enhancer
of split grouch protein
Eur. J. Biochem. 216 (1), 343-352 (1993)
93373944

2 (bases 1 to 12524)
Hou, E.W. and Li, S.S.-L.
REFERENCE Genomic organization and in situ localization to chromosome 19p13.3
AUTHORS of human AES gene: the gene product exhibiting strong similarity to
TITLE the amino-terminal domain of Drosophila enhancer of split grouch
protein
Unpublished

JOURNAL 3 (bases 1 to 12524)
REFERENCE Hou, E.W. and Li, S.S.-L.
AUTHORS Direct Submission

TITLE Submitted (06-FEB-1997) Laboratory of Molecular Genetics, National
JOURNAL Institute of Environmental Health Science, Mail Drop D3-05, P.O.
Box 12233, Research Triangle Park, NC 27709, USA

FEATURES Location/Qualifiers

source

1. .12524
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="19p13.3"

/tissue_type="lymphocyte and umbilical cord"

mrna join(U88831.1:1..57,3046..3143,6662..6725,8436..8470,
9055..9117,10921..10995,11075..>11296)
/gene="AES"

/product="groucho protein homolog"

gene join(U88831.1:1..59,1..11296)

/gene="AES"

CDS join(U88831.1:31..57,3046..3143,6662..6725,8436..8470,
9055..9117,10921..10995,11075..11296)
/gene="AES"

/note="similar to human grouch protein encoded by GenBank
Accession Number U04241 and to Drosophila melanogaster
grouch protein: Swiss Prot Accession Number P16371; the
corresponding mRNA sequence is reported in GenBank
Accession Numbers X73357 and X73358"

/codon_start=1

/product="grouch protein homolog"

protein_id="AAD00654.1"

/db_xref="GI:4099601"

/translation="MMFPQSRHSGSHLPQOLKFTTSDSCDRITDFQLQAQYHSLK
LECKIASEKEMQRHYVYVYEMSYGLNIEMHKQAEIVKRLNGICAVLPYLSQHQHQ
QVIGAIERAKQVTAPELNSITRQLOAHQLSQLOALALPLTLPVGLQPPSLPAVSAG
TGLLSLSALGSQLSKEDKNHGDDTHQEDDGKSD"
1..3045

/gene="AES"

/number=1

3046..3143

/gene="AES"

/number=2

3144..6661

/gene="AES"

/number=2

6662..6725

/gene="AES"

/number=3

6726..8425

/gene="AES"

/number=3

8426..8470

/gene="AES"

/number=4

8471..9054

/gene="AES"

/number=4

9055..9117

/gene="AES"

/number=5

9118..10920

/gene="AES"

/number=5

10921..10995

/gene="AES"

/number=6

10996..11074

/gene="AES"

/number=6

11075..>11296

/gene="AES"

/number=7

polyA_signal 11923..11928

polyA_signal 12173..12178

BASE COUNT 2877 a 3347 c 3480 g 2811 t

ORIGIN 9 others

Query Match 1.9%; Score 20; DB 9; Length 12524;

Best Local Similarity 100.0%; Pred. No. 31;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 857 CCAATTTCTACACAAAGCTC 876

|||||

Db 5951 CCAATTTCTACACAAAGCTC 5970

AC002997 37784 bp DNA linear PRI 07-OCT-1997

Human DNA from chromosome 19-specific cosmid R46667, genomic

sequence, complete sequence.

AC002997

AC002997

AC002997.1 GI:2477519

HTG.

human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 37784)

AUTHORS Lamerdin, J.E., McCreedy, P.M., Adamson, A.W., Burkhardt-Schultz, K.,

Garcia, E., Kyle, A., Ramirez, M., Stilwagen, S., Garnes, J.,

Danganan, L., Bruce, R., Quan, G., Montgomery, M., Ow, D., Kobayashi, A.,

Olsen, A.O. and Carrano, A.V.

Sequence analysis of a 1 Mb region in 19q13.1

Unpublished

TITLE

JOURNAL

REFERENCE	2 (bases 1 to 37784)
AUTHORS	Lamerdin, J.E.
TITLE	Direct Submission
JOURNAL	Submitted (07-OCT-1997) Human Genome Center, Lawrence Livermore National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
COMMENT	Map and sequence orientated from centromere to telomere. R2667 overlaps PAC PC28130 to the left and cosmid F21426 to the right.
FEATURES	<p>Location/Qualifiers</p> <p>1. .37784</p> <p>/organism="Homo sapiens"</p> <p>/db_xref="taxon:9606"</p> <p>/chromosome="19"</p> <p>/map="19q13.1 from D19S208 to COX7A1"</p> <p>/clone="R26667"</p> <p>/cell_line="5HL2-B"</p> <p>/clone_lib="LL19NC03 R chromosome 19-specific cosmid library"</p> <p>/note="LL19NC03 library constructed at LLNL from flow-sorted chromosomes from hybrid 5HL2-B, which carries chromosome 19 as its only human chromosome."</p> <p>4. .178</p> <p>/rpt_family="Alu"</p> <p>complement(788. .1056)</p> <p>/rpt_family="Alu"</p> <p>2897. .3360</p> <p>/rpt_family="Alu"</p> <p>complement(3631. .3716)</p> <p>/rpt_family="MLT1"</p> <p>complement(4023. .4206)</p> <p>/note="predicted exon, program: graill2exons_human_1.3, frame: 1, quality: marginal, score: 49.000"</p> <p>complement(4383. .4678)</p> <p>/rpt_family="Alu"</p> <p>complement(4690. .4830)</p> <p>/rpt_family="MLT1"</p> <p>5448. .5858</p> <p>/rpt_family="Alu"</p> <p>6150. .6602</p> <p>/rpt_family="Alu"</p> <p>7409. .7685</p> <p>/rpt_family="Alu"</p> <p>8018. .8305</p> <p>/rpt_family="Alu"</p> <p>complement(8423. .8722)</p> <p>/rpt_family="Alu"</p> <p>8762. .8861</p> <p>/rpt_family="LI"</p> <p>8866. .9141</p> <p>/rpt_family="Alu"</p> <p>9164. .9267</p> <p>/rpt_family="LI"</p> <p>complement(9590. .9869)</p> <p>/rpt_family="Alu"</p> <p>9988. .10088</p> <p>/note="predicted exon, program: graill2exons_human_1.3, frame: 2, quality: excellent, score: 85.000"</p> <p>10463. .10759</p> <p>/rpt_family="Alu"</p> <p>11550. .11683</p> <p>/note="predicted exon, program: graill2exons_human_1.3, frame: 2, quality: good, score: 64.000"</p> <p>12808. .13086</p> <p>/rpt_family="Alu"</p> <p>complement(13957. .14248)</p> <p>/rpt_family="Alu"</p> <p>14871. .15863</p> <p>/note="predicted exon, program: graill2exons_human_1.3, frame: 2, quality: excellent, score: 92.000"</p> <p>16106. .16395</p> <p>/rpt_family="Alu"</p> <p>17263. .17693</p> <p>/rpt_family="Alu"</p>
repeat_region	complement(18158. .18401)
repeat_region	/rpt_family="Alu"
repeat_region	18807. .19120
repeat_region	/rpt_family="Alu"
repeat_region	19298. .19574
repeat_region	/rpt_family="Alu"
repeat_region	complement(20158. .20442)
repeat_region	/rpt_family="Alu"
repeat_region	20491. .20791
repeat_region	/rpt_family="Alu"
repeat_region	complement(21064. .21217)
repeat_region	/rpt_family="Alu"
repeat_region	complement(21492. .21693)
repeat_region	/rpt_family="Alu"
repeat_region	21695. .22143
repeat_region	/rpt_family="Alu"
repeat_region	22970. .23020
repeat_region	/rpt_family="MER5"
misc_feature	complement(24589. .24894)
misc_feature	/note="DDS similarity to T03155 PB26A8 Fetal brain, Stratagene Homo sapiens cDNA clone PB26A8 3'end (1. .327). Score: 581 Identity: 313/327 (95%)"
repeat_region	25560. .25849
repeat_region	/rpt_family="Alu"
repeat_region	complement(26656. .26942)
repeat_region	/rpt_family="Alu"
repeat_region	26991. .27346
repeat_region	/rpt_family="Alu"
misc_feature	complement(28264. .28323)
misc_feature	/note="predicted exon, program: graill2exons_human_1.3, frame: 1, quality: good, score: 74.000"
repeat_region	complement(28345. .28635)
repeat_region	/rpt_family="Alu"
misc_feature	complement(28730. .28894)
misc_feature	/note="predicted exon, program: graill2exons_human_1.3, frame: 0, quality: good, score: 61.000"
repeat_region	complement(28907. .29188)
repeat_region	/rpt_family="Alu"
repeat_region	29333. .29428
repeat_region	/rpt_family="LTR10"
repeat_region	complement(29828. .30108)
repeat_region	/rpt_family="Alu"
repeat_region	30963. .30998
repeat_region	/rpt_family="Alu"
repeat_region	complement(31083. .31369)
repeat_region	/rpt_family="Alu"
repeat_region	31375. .31539
repeat_region	/rpt_family="Alu"
repeat_region	complement(32052. .32709)
repeat_region	/rpt_family="Alu"
repeat_region	33521. .33758
repeat_region	/rpt_family="Alu"
repeat_region	complement(34242. .34537)
repeat_region	/rpt_family="Alu"
repeat_region	34671. .34879
repeat_region	/rpt_family="MER7"
repeat_region	34927. .34965
repeat_region	/rpt_family="MER7"
repeat_region	35188. .35456
repeat_region	/rpt_family="Alu"
repeat_region	35545. .35669
misc_feature	/note="predicted exon, program: graill2exons_human_1.3, frame: 0, quality: excellent, score: 85.000"
repeat_region	complement(35688. .35869)
repeat_region	/rpt_family="Alu"
repeat_region	36471. .36532
repeat_region	/rpt_family="MER4"
repeat_region	36565. .36855
repeat_region	/rpt_family="Alu"
repeat_region	36861. .36994
repeat_region	/rpt_family="MER4"
repeat_region	complement(37144. .37245)
repeat_region	/rpt_family="MIR"

BASE COUNT 10418 a 8913 c 8981 g 9472 t

Query Match 1.9%; Score 20; DB 9; Length 37784;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 189 CCTTATGATCTGCCTGCTT 208
|||||
Db 21920 CCTTATGATCTGCCTGCTT 21901

RESULT 45

AC105178
LOCUS Homo sapiens chromosome 8 clone RP11-110I16 map 8, LOW-PASS
DEFINITION AC105178
ACCESSION AC105178
VERSION 1
KEYWORDS HTG; HTGS_PHASE0.
SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 63400)

REFERENCE

AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE

JOURNAL Homo sapiens chromosome 8, clone RP11-110I16

AUTHORS

1 (bases 1 to 63400)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bouckghalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Katat,A., Karatas,A., Kells,C., LaRoque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mienna,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE

JOURNAL Submitted (26-DEC-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L21931

Center clone name: 110_I_16

* NOTE: This record contains 79 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.

* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 693: contig of 693 bp in length
694 793: gap of 100 bp
794 1514: contig of 721 bp in length
1515 1614: gap of 100 bp
1615 2322: contig of 708 bp in length
2323 2422: gap of 100 bp
2423 3109: contig of 687 bp in length
3110 3209: gap of 100 bp
3210 3905: contig of 696 bp in length
3906 4005: gap of 100 bp
4006 4711: contig of 706 bp in length
4712 4811: gap of 100 bp
4812 5490: contig of 679 bp in length
5491 5590: gap of 100 bp
5591 6268: contig of 678 bp in length
6269 6368: gap of 100 bp
6369 7066: contig of 698 bp in length
7067 7166: gap of 100 bp
7167 7868: contig of 702 bp in length
7869 7968: gap of 100 bp
7969 8677: contig of 709 bp in length
8678 8777: gap of 100 bp
8778 9431: contig of 654 bp in length
9432 9531: gap of 100 bp
9532 10241: contig of 710 bp in length
10242 10341: gap of 100 bp
10342 11060: contig of 719 bp in length
11061 11160: gap of 100 bp
11161 11856: contig of 696 bp in length
11857 11956: gap of 100 bp
11957 12657: contig of 701 bp in length
12658 12757: gap of 100 bp
12758 13467: contig of 710 bp in length
13468 13567: gap of 100 bp
13568 14275: contig of 708 bp in length
14276 14375: gap of 100 bp
14376 15082: contig of 707 bp in length
15083 15182: gap of 100 bp
15183 15862: contig of 680 bp in length
15863 15962: gap of 100 bp
15963 16681: contig of 719 bp in length
16682 16781: gap of 100 bp
16782 17495: contig of 714 bp in length
17496 17595: gap of 100 bp
17596 18286: contig of 691 bp in length
18287 18386: gap of 100 bp
18387 19103: contig of 717 bp in length
19104 19203: gap of 100 bp
19204 19916: contig of 713 bp in length
19917 20016: gap of 100 bp
20017 20715: contig of 699 bp in length
20716 20815: gap of 100 bp
20816 21510: contig of 695 bp in length
21511 21610: gap of 100 bp
21611 22338: contig of 728 bp in length
22339 22438: gap of 100 bp
22439 23136: contig of 698 bp in length
23137 23236: gap of 100 bp
23237 23951: contig of 715 bp in length
23952 24051: gap of 100 bp
24052 24756: contig of 705 bp in length
24757 24856: gap of 100 bp
24857 25556: contig of 700 bp in length
25557 25656: gap of 100 bp
25657 26370: contig of 714 bp in length
26371 26470: gap of 100 bp
26471 27184: contig of 714 bp in length
27185 27284: gap of 100 bp
27285 27985: contig of 701 bp in length

* 27986 28085: gap of 100 bp
* 28086 28793: contig of 708 bp in length
* 28794 28893: gap of 100 bp
* 28894 29592: contig of 699 bp in length
* 29593 29692: gap of 100 bp
* 29693 30416: contig of 724 bp in length
* 30417 30516: gap of 100 bp
* 30517 31200: contig of 684 bp in length
* 31201 31300: gap of 100 bp
* 31301 32009: contig of 709 bp in length
* 32010 32109: gap of 100 bp
* 32110 32815: contig of 706 bp in length
* 32816 32915: gap of 100 bp
* 32916 33623: contig of 708 bp in length
* 33624 33723: gap of 100 bp
* 33724 34430: contig of 707 bp in length
* 34431 34530: gap of 100 bp
* 34531 35248: contig of 718 bp in length
* 35249 35348: gap of 100 bp
* 35349 36056: contig of 708 bp in length
* 36057 36156: gap of 100 bp
* 36157 36864: contig of 708 bp in length
* 36865 36964: gap of 100 bp
* 36965 37676: contig of 712 bp in length
* 37677 37776: gap of 100 bp
* 37777 38482: contig of 706 bp in length
* 38483 38582: gap of 100 bp
* 38583 39299: contig of 717 bp in length
* 39300 39399: gap of 100 bp
* 39400 40095: contig of 696 bp in length
* 40096 40195: gap of 100 bp
* 40196 40915: contig of 720 bp in length
* 40916 41015: gap of 100 bp
* 41016 41710: contig of 695 bp in length
* 41711 41810: gap of 100 bp
* 41811 42516: contig of 706 bp in length
* 42517 42616: gap of 100 bp
* 42617 43311: contig of 695 bp in length
* 43312 43411: gap of 100 bp
* 43412 44122: contig of 711 bp in length
* 44123 44222: gap of 100 bp
* 44223 44941: contig of 719 bp in length
* 44942 45041: gap of 100 bp
* 45042 45767: contig of 726 bp in length
* 45768 45867: gap of 100 bp
* 45868 46577: contig of 710 bp in length
* 46578 46677: gap of 100 bp
* 46678 47379: contig of 702 bp in length
* 47380 47479: gap of 100 bp
* 47480 48180: contig of 701 bp in length
* 48181 48280: gap of 100 bp
* 48281 48978: contig of 698 bp in length
* 48979 49078: gap of 100 bp
* 49079 49764: contig of 686 bp in length
* 49765 49864: gap of 100 bp
* 49865 50572: contig of 708 bp in length
* 50573 50672: gap of 100 bp
* 50673 51345: contig of 673 bp in length
* 51346 51445: gap of 100 bp
* 51446 52152: contig of 707 bp in length
* 52153 52252: gap of 100 bp
* 52253 52946: contig of 694 bp in length
* 52947 53046: gap of 100 bp
* 53047 53751: contig of 705 bp in length
* 53752 53851: gap of 100 bp
* 53852 54565: contig of 714 bp in length
* 54566 54665: gap of 100 bp
* 54666 55376: contig of 711 bp in length

Query Match 1.9%: Score 20; DB 2; Length 63400;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 189 CCTATGATCTGCCTT 208
| | | | | | | | | | | | | | | | | | | | | |
Db 8204 CCTATGATCTGCCTT 8223

Search completed: October 29, 2002, 04:47:09
Job time : 3464 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2002, 02:43:07 ; Search time 48 Seconds
(without alignments)
5327.177 Million cell updates/sec

Title: US-09-886-041-1
Perfect score: 1041
Sequence: 1 atgtacaacgggtcgtgctg.....acattgttgaggcactga 1041

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0
Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*1: /cgn2.6/ptodata/1/ina/5A_COMB.seq:*2: /cgn2.6/ptodata/1/ina/5B_COMB.seq:*3: /cgn2.6/ptodata/1/ina/6A_COMB.seq:*4: /cgn2.6/ptodata/1/ina/6B_COMB.seq:*5: /cgn2.6/ptodata/1/ina/PTUS_COMB.seq:*6: /cgn2.6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18	1.7	433	1 US-08-664-596B-1	Sequence 1, Appl
C 2	18	1.7	433	1 US-08-738-367-1	Sequence 1, Appl
C 3	18	1.7	733	1 US-08-738-367-6	Sequence 6, Appl
C 4	18	1.7	3380	2 US-09-156-425-1	Sequence 1, Appl
C 5	18	1.7	6453	1 US-08-306-691B-14	Sequence 14, Appl
C 6	18	1.7	6453	3 US-09-209-668-10	Sequence 10, Appl
C 7	18	1.7	6453	3 US-09-356-952-8	Sequence 8, Appl
C 8	18	1.7	7898	4 US-08-984-709A-49	Sequence 49, Appl
C 9	17	1.6	1125	3 US-07-928-462-1	Sequence 1, Appl
C 10	17	1.6	1125	3 US-08-273-247-1	Sequence 1, Appl
C 11	17	1.6	1935	4 US-09-423-890-5	Sequence 5, Appl
C 12	17	1.6	3264	2 US-08-708-541A-31	Sequence 31, Appl
C 13	17	1.6	3264	2 US-08-708-541A-33	Sequence 31, Appl
C 14	17	1.6	4138	1 US-08-447-411-75	Sequence 75, Appl
C 15	17	1.6	4138	2 US-08-662-227-33	Sequence 33, Appl
C 16	17	1.6	4138	4 US-09-017-947-33	Sequence 33, Appl
C 17	17	1.6	6822	4 US-09-426-998-3	Sequence 3, Appl
C 18	17	1.6	7741	4 US-09-426-998-4	Sequence 4, Appl
C 19	16	1.5	20	4 US-09-487-368A-226	Sequence 226, App
C 20	16	1.5	30	1 US-08-484-557C-10	Sequence 10, Appl
C 21	16	1.5	30	1 US-08-487-426B-10	Sequence 10, Appl
C 22	16	1.5	30	2 US-08-487-720A-10	Sequence 10, Appl
C 23	16	1.5	78	3 US-08-945-734-10	Sequence 10, Appl
C 24	16	1.5	78	4 US-09-258-797-10	Sequence 10, Appl
C 25	16	1.5	78	5 PCT-US96-09451-10	Sequence 10, Appl
C 26	16	1.5	99	4 US-09-060-756-531	Sequence 531, App
C 27	16	1.5	287	4 US-09-060-756-386	Sequence 386, App

28	1.5	315	4	US-09-060-756-46	Sequence 46, Appl
C 29	16	1.5	335	4 US-08-576-202-30	Sequence 30, Appl
C 30	16	1.5	335	5 PCT-US95-16766-30	Sequence 30, Appl
C 31	16	1.5	440	2 US-08-967-101-122	Sequence 122, App
C 32	16	1.5	440	2 US-08-592-541-122	Sequence 122, App
C 33	16	1.5	440	3 US-09-124-698-122	Sequence 122, App
C 34	16	1.5	440	4 US-09-127-480-122	Sequence 122, App
C 35	16	1.5	440	4 US-08-496-841C-122	Sequence 122, App
C 36	16	1.5	450	2 US-08-967-101-7	Sequence 7, Appl
C 37	16	1.5	450	2 US-08-967-101-73	Sequence 73, Appl
C 38	16	1.5	450	2 US-08-592-541-7	Sequence 7, Appl
C 39	16	1.5	450	2 US-08-592-541-73	Sequence 73, Appl
C 40	16	1.5	450	3 US-09-124-698-7	Sequence 7, Appl
C 41	16	1.5	450	3 US-09-124-698-73	Sequence 73, Appl
C 42	16	1.5	450	4 US-09-127-480-7	Sequence 7, Appl
C 43	16	1.5	450	4 US-09-127-480-73	Sequence 73, Appl
C 44	16	1.5	450	4 US-08-496-841C-7	Sequence 7, Appl
C 45	16	1.5	450	4 US-08-496-841C-73	Sequence 73, Appl

ALIGNMENTS

RESULT 1
US-08-664-596B-1/c
; Sequence 1, Application US/08664596B
; Patent No. 5807703
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: LaVallie, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Evans, Cheryl
; APPLICANT: Spaulding, Vikki
; APPLICANT: Bowman, Michael
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/664,596B
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 433 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-664-596B-1

Query Match 1.7%; Score 18; DB 1; Length 433;
Best Local Similarity 100.0%; Pred. No. 11;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 531 TGACATCATGTTCCAGCT 548
|||||
Db 341 TGACATCATGTTCCAGCT 324

RESULT 2

US-08-738-367-1/c
; Sequence 1, Application US/08738367
; Patent No. 5827688
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: Lavallie, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/738,367
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 433 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-738-367-1

Query Match 1.7%; Score 18; DB 1; Length 433;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 531 TGACATCATGTTCCAGCT 548
|||||
Db 341 TGACATCATGTTCCAGCT 324

RESULT 3

US-08-738-367-6/c
; Sequence 6, Application US/08738367
; Patent No. 5827688
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: Lavallie, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice

; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/738,367
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 733 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-738-367-6

Query Match 1.7%; Score 18; DB 1; Length 733;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 531 TGACATCATGTTCCAGCT 548
|||||
Db 341 TGACATCATGTTCCAGCT 324

RESULT 4

US-09-156-425-1/c
; Sequence 1, Application US/09156425B
; Patent No. 5962671
; GENERAL INFORMATION:
; APPLICANT: Baker, Brenda F.
; APPLICANT: Cowsett, Lex M.
; TITLE OF INVENTION: ANTISENSE MODULATION OF FAN EXPRESSION
; FILE REFERENCE: PFS-0009
; CURRENT APPLICATION NUMBER: US/09/156,425B
; CURRENT FILING DATE: 1998-09-18
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
; LENGTH: 3380
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)..(2766)
US-09-156-425-1

Query Match 1.7%; Score 18; DB 2; Length 3380;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 531 TGACATCATGTTCCAGCT 548
|||||
Db 2305 TGACATCATGTTCCAGCT 2288

```
RESULT 5
US-08-306-691B-14/c
; Sequence 14, Application US/08306691B
; Patent No. 5734039
; GENERAL INFORMATION:
; APPLICANT: Calabretta, Bruno
; APPLICANT: Skorski, Tomasz
; TITLE OF INVENTION: ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Sedel, Gonda, Lavorgna & Monaco, P.C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306,691B
; FILING DATE: September 15, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5349
; TELEX: No. 5734039e
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6453 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-306-691B-14

Query Match 1.7%; Score 18; DB 1; Length 6453;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 188 TCCTTATGATCTGCCCTGC 205
|||||
Db 3966 TCCTTATGATCTGCCCTGC 3949

RESULT 6
US-09-209-668-10/c
; Sequence 10, Application US/09209668A
; Patent No. 6114517
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Xu, Xiaoxing S.
; TITLE OF INVENTION: METHODS OF MODULATING TUMOR NECROSIS FACTOR
; TITLE OF INVENTION: alpha-INDUCED EXPRESSION OF CELL ADHESION MOLECULES
; FILE REFERENCE: ISPH-0336
; CURRENT APPLICATION NUMBER: US/09/209,668A
; CURRENT FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 6453
; TYPE: DNA
; ORGANISM: Homo sapiens

Query Match 1.7%; Score 18; DB 1; Length 6453;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 188 TCCTTATGATCTGCCCTGC 205
|||||
Db 3966 TCCTTATGATCTGCCCTGC 3949

RESULT 7
US-09-356-952-8/c
; Sequence 8, Application US/09356952
; Patent No. 6117663
; GENERAL INFORMATION:
; APPLICANT: Boriack-Sjodin, Ann
; APPLICANT: Margarit, S. M.
; APPLICANT: Bor-Soghl, Dafna
; APPLICANT: Cole, Philip
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: A CRYSTAL OF A RAS-SOS COMPLEX AND METHODS OF USE
; FILE REFERENCE: 600-1-228N
; CURRENT APPLICATION NUMBER: US/09/356,952
; CURRENT FILING DATE: 1999-07-19
; EARLIER APPLICATION NUMBER: 60/093,631
; EARLIER FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 6453
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-356-952-8

Query Match 1.7%; Score 18; DB 3; Length 6453;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 188 TCCTTATGATCTGCCCTGC 205
|||||
Db 3966 TCCTTATGATCTGCCCTGC 3949

RESULT 8
US-08-984-709A-49/c
; Sequence 49, Application US/08984709A
; Patent No. 6320032
; GENERAL INFORMATION:
; APPLICANT: Williams, Mark E.
; APPLICANT: Stauderman, Kenneth A.
; APPLICANT: Harpold, Michael M.
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe
```

```
Query Match 1.7%; Score 18; DB 3; Length 6453;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 188 TCCTTATGATCTGCCCTGC 205
|||||
Db 3966 TCCTTATGATCTGCCCTGC 3949
```

STREET: 4250 Executive Square, Suite 700
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/984,709A
FILING DATE: 02-DEC-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24735-9815 (formerly 6362-9815)
TELEPHONE: (619) 450-8400
TELEFAX: (619) 587-5360
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 7898 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 249...7307
OTHER INFORMATION:
US-08-984-709A-49

Query Match 1.7%; Score 18; DB 4; Length 7898;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 610 AGCGGGAGCGACGACGTG 627
|||||

Db 4009 AGCGGGAGCGACGACGTG 3992

RESULT 9
US-07-928-462-1/c
Sequence 1, Application US/07928462
Patent No. 5328996
GENERAL INFORMATION:
APPLICANT: Boyle, Michael D.P.
APPLICANT: Lottenberg, Richard
APPLICANT: Broder, Christopher C.
APPLICANT: von Mering, Gregory O.
TITLE OF INVENTION: Bacterial Plasmin Receptors as
TITLE OF INVENTION: Fibrinolytic Agents
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/928,462

FILING DATE: 19920810
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/524,411
FILING DATE: 16-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/330,849
FILING DATE: 29-MAR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF/S&S-13.C2
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1125 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: both
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Streptococcus pyogenes
STRAIN: M untypeable
INDIVIDUAL ISOLATE: 64/14
IMMEDIATE SOURCE:
CLONE: pRL015
FEATURE:
NAME/KEY: CDS
LOCATION: 115..1122
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 115..1122
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /codon_start= 115
OTHER INFORMATION: /function= "High-affinity binding of
OTHER INFORMATION: plasmin(ogen)"
OTHER INFORMATION: /product= "Streptococcal plasmin receptor"
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "plr"
OTHER INFORMATION: /number= 1
OTHER INFORMATION: /label= PLR
US-07-928-462-1

Query Match 1.6%; Score 17; DB 1; Length 1125;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 180 TGATTTCCTCTTATGA 196

Db 111 TGATTTCCTCTTATGA 95

RESULT 10
US-08-273-247-1/c
Sequence 1, Application US/08273247
Patent No. 6136323
GENERAL INFORMATION:
APPLICANT: Boyle, Michael D.P.
APPLICANT: Lottenberg, Richard
APPLICANT: Broder, Christopher C.
APPLICANT: von Mering, Gregory O.
TITLE OF INVENTION: Bacterial Plasmin Receptors as
TITLE OF INVENTION: Fibrinolytic Agents
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL

```
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/273,247
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/524,411
; FILING DATE: 16-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/330,849
; FILING DATE: 29-MAR-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF/S&S-13.C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1125 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes
; STRAIN: M untypable
; INDIVIDUAL ISOLATE: 64/14
; IMMEDIATE SOURCE:
; CLONE: pRL015
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 115..1122
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 115..1122
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /codon_start=115
; OTHER INFORMATION: /function= "High-affinity binding of
; OTHER INFORMATION: plasmin(ogen)"
; OTHER INFORMATION: /product= "Streptococcal plasmin receptor"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "Plr"
; OTHER INFORMATION: /number= 1
; OTHER INFORMATION: /label= PLR
;
; US-08-273-247-1
;
; Query Match 1.6%; Score 17; DB 3; Length 1125;
; Best Local Similarity 100.0%; Pred. No. 37;
; Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 180 TGATTTCTCTCTTATGA 196
; |||||
; DB 111 TGATTTCTCTCTTATGA 95
;
; RESULT 11
; US-09-423-890-5
; Sequence 5, Application US/09423890
; Patent No. 6312934
; GENERAL INFORMATION:
; APPLICANT: CADUS PHARMACEUTICAL CORPORATION
; TITLE OF INVENTION: HUMAN MERK PROTEIN AND NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: AND USES THEREFOR
```

```
; FILE REFERENCE: CPI-085CPPC
; CURRENT APPLICATION NUMBER: US/09/423,890
; CURRENT FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: USN 60/078,153
; PRIOR FILING DATE: 1998-03-16
; PRIOR APPLICATION NUMBER: USN 60/099,165
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1935
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25)..(1902)
; US-09-423-890-5
;
; Query Match 1.6%; Score 17; DB 4; Length 1935;
; Best Local Similarity 100.0%; Pred. No. 38;
; Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 812 GCATGCTGGATCCCTG 828
; |||||
; DB 659 GCATGCTGGATCCCTG 675
;
; RESULT 12
; US-08-708-541A-31/c
; Sequence 31, Application US/08708541A
; Patent No. 5871744
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, Vikram N.
; APPLICANT: MUNDT, Egbert
; TITLE OF INVENTION: A METHOD FOR GENERATING BIRNAVIRUS FROM
; TITLE OF INVENTION: SYNTHETIC RNA TRANSCRIPTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKAIIDO, MARCELSTEIN, MURRAY & ORAM LLP
; STREET: 655 Fifteenth Street, N. W.,
; STREET: Suite 330 - G Street Lobby
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/708,541A
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: KITTS, Monica C.
; REGISTRATION NUMBER: 36,105
; REFERENCE/DOCKET NUMBER: P8172-6002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202/638-5000
; TELEFAX: 202/638-4810
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 97..531
; US-08-708-541A-31
```

Query Match 1.6%; Score 17; DB 2; Length 3264;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 CCACATGAAGACCTGGA 142
|||||
Db 2506 CCACATGAAGACCTGGA 2490

RESULT 13
US-08-708-541A-33/c
; Sequence 33, Application US/08708541A
; Patent No. 5871744
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, Vikram N.
; APPLICANT: MUNDT, Egbert
; TITLE OF INVENTION: A METHOD FOR GENERATING BIRNAVIRUS FROM
; TITLE OF INVENTION: SYNTHETIC RNA TRANSCRIPTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKATDO, MARCELSTEIN, MURRAY & ORAM LLP
; STREET: 655 Fifteenth Street, N. W.,
; STREET: Suite 330 - G Street Lobby
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/708,541A
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: KITTS, Monica C.
; REGISTRATION NUMBER: 36,105
; REFERENCE/DOCKET NUMBER: P8172-6002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202/638-5000
; TELEFAX: 202/638-4810
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 131...3169
US-08-708-541A-33

Query Match 1.6%; Score 17; DB 2; Length 3264;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 CCACATGAAGACCTGGA 142
|||||
Db 2506 CCACATGAAGACCTGGA 2490

RESULT 14
US-08-447-411-75
; Sequence 75, Application US/08447411
; Patent No. 5773243
; GENERAL INFORMATION:
; APPLICANT: FRITZINGER, DAVID C.
; APPLICANT: BREDEHORST, REINHARD
; APPLICANT: VOGEL, CARL-WILHELM
; TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2

; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,411
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/043,747
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5773243man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-101-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 75:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4138 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..4001
US-08-447-411-75

Query Match 1.6%; Score 17; DB 1; Length 4138;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 870 CAAGCTCAAAATCTGCA 886
|||||
Db 3149 CAAGCTCAAAATCTGCA 3165

RESULT 15
US-08-662-227-33
; Sequence 33, Application US/08662227
; Patent No. 5922320
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; APPLICANT: BREDEHORST, REINHORST
; APPLICANT: KOCK, MICHAEL
; APPLICANT: FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/662,227
; FILING DATE: 14-JUN-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-0107-0X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4138 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-662-227-33

Query Match 1.6%; Score 17; DB 2; Length 4138;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 870 CAAGCTCAAAATCTGCA 886
|||||
Db 3149 CAAGCTCAAAATCTGCA 3165

RESULT 16
US-09-017-947-33
; Sequence 33, Application US/09017947
; Patent No. 6303754
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; APPLICANT: BREDEHORST, REINHORST
; APPLICANT: KOCK, MICHAEL
; APPLICANT: FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUBSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/017,947
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/662,227
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-0107-0X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4138 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-017-947-33

Query Match 1.6%; Score 17; DB 4; Length 4138;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 870 CAAGCTCAAAATCTGCA 886
|||||
Db 3149 CAAGCTCAAAATCTGCA 3165

RESULT 17
US-09-426-998-3
; Sequence 3, Application US/09426998
; Patent No. 6358706
; GENERAL INFORMATION:
; APPLICANT: DUBIN, ADRIENNE E.
; APPLICANT: PYATI, JAYASHREE
; APPLICANT: ZHU, JESSICA Y
; APPLICANT: ERLANDER, MARK G
; APPLICANT: GALINDO, JOSE E
; TITLE OF INVENTION: DNA ENCODING HUMAN ALPHA1G T-TYPE CALCIUM
; FILE REFERENCE: ORT-1057
; CURRENT APPLICATION NUMBER: US/09/426,998
; CURRENT FILING DATE: 1999-10-26
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PATENTIN VER. 2.0
; SEQ ID NO 3
; LENGTH: 6822
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
; US-09-426-998-3

Query Match 1.6%; Score 17; DB 4; Length 6822;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 CTGCTCATTGTGGCCTT 77
|||||
Db 4561 CTGCTCATTGTGGCCTT 4577

RESULT 18
US-09-426-998-4
; Sequence 4, Application US/09426998
; Patent No. 6358706
; GENERAL INFORMATION:
; APPLICANT: DUBIN, ADRIENNE E.
; APPLICANT: PYATI, JAYASHREE
; APPLICANT: ZHU, JESSICA Y
; APPLICANT: ERLANDER, MARK G
; APPLICANT: GALINDO, JOSE E
; TITLE OF INVENTION: DNA ENCODING HUMAN ALPHA1G T-TYPE CALCIUM
; FILE REFERENCE: ORT-1057
; CURRENT APPLICATION NUMBER: US/09/426,998
; CURRENT FILING DATE: 1999-10-26
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PATENTIN VER. 2.0
; SEQ ID NO 4
; LENGTH: 7741
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
; US-09-426-998-4

Query Match 1.6%; Score 17; DB 4; Length 7741;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 CTGCTCATTGTGGCCTT 77
```

Db 5083 CTGCTCATGTGGCCTT 5099
|||||

RESULT 19
US-09-487-368A-226/c
; Sequence 226, Application US/09487368A
; Patent No. 6261840
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowser
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF PTP1B EXPRESSION
; FILE REFERENCE: RTS-0093
; CURRENT APPLICATION NUMBER: US/09/487,368A
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 240
; SEQ ID NO 226
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-09-487-368A-226

Query Match 1.5%; Score 16; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 431 TCCTGGGAACAGTGA 446
Db 16 TCCTGGGAACAGTGA 1
|||||

RESULT 20
US-08-484-557C-10
; Sequence 10, Application US/08484557C
; Patent No. 5693502
; GENERAL INFORMATION:
; APPLICANT: LARRY GOLD
; APPLICANT: SUMEDHA JAYASENA
; TITLE OF INVENTION: NUCLEIC ACID LIGAND
; TITLE OF INVENTION: INHIBITORS TO DNA POLYMERASES
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson and Bratschun, L.L.C.
; STREET: 8400 East Prentice Avenue., Suite 200
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb
; MEDIUM TYPE: Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,557C
; FILING DATE: 7-JUNE-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/714,131
; FILING DATE: 10-JUNE-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/536,428
; FILING DATE: 11-JUNE-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/964,624
; FILING DATE: 21-OCTOBER-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Diane Cruz
; REGISTRATION NUMBER: 33,960
; REFERENCE/DOCKET NUMBER: NEX43-3
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3333
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-484-557C-10

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-484-557C-10

Query Match 1.5%; Score 16; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 GCTTTGGGACATTC 259
Db 2 GCTTTGGGACATTC 17
|||||

RESULT 21
US-08-487-426B-10
; Sequence 10, Application US/08487426B
; Patent No. 5763173
; GENERAL INFORMATION:
; APPLICANT: LARRY GOLD
; APPLICANT: SUMEDHA JAYASENA
; TITLE OF INVENTION: NUCLEIC ACID LIGAND
; TITLE OF INVENTION: INHIBITORS TO DNA POLYMERASES
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson and Bratschun, L.L.C.
; STREET: 8400 East Prentice Avenue., Suite 200
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb
; MEDIUM TYPE: Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,426B
; FILING DATE: 7-JUNE-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/714,131
; FILING DATE: 10-JUNE-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/536,428
; FILING DATE: 11-JUNE-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/964,624
; FILING DATE: 21-OCTOBER-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Diane Cruz
; REGISTRATION NUMBER: 33,960
; REFERENCE/DOCKET NUMBER: NEX43-1
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-487-426B-10

Query Match 1.5%; Score 16; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels

QY 244 GCTTTTGGGGACATTC 259
Db 2 GCTTTTGGGGACATTC 17

RESULT 22
US-08-487-720A-10
; Sequence 10, Application US/08487720A
; Patent No. 5874557
; GENERAL INFORMATION:
; APPLICANT: LARRY GOLD
; APPLICANT: SOMEDHA JAYASENA
; TITLE OF INVENTION: NUCLEIC ACID LIGAND
; TITLE OF INVENTION: INHIBITORS TO DNA POLYMERASES
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson and Bratschun, L.L.C.
; STREET: 8400 East Prentice Avenue, Suite 200
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111

Query Match 1.5%; Score 16; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. NO. 1.1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels

Qy	244	GCTTTTGGGGACATTC	259
Db	2	GCTTTTGGGGACATTC	17

RESULT 23
US-08-945-734-10
; Sequence 10, Application US/08945734

Patent No. 6020130
GENERAL INFORMATION:
APPLICANT: LARRY GOLD
APPLICANT: SUMEDHA JAYASENA
TITLE OF INVENTION: NUCLEIC ACID LIGANDS
TITLE OF INVENTION: THAT BIND TO AND INHIBIT DNA
TITLE OF INVENTION: POLYMERASES
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson and Bratschun, L.L.C.
STREET: 8400 East Prentice Avenue., Suite 200
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80111

```

Query Match          1.5%; Score 16; DB 3; Length 78;
Best Local Similarity 100.0%; Pred. No. 1.le+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 244 GCTTTTGGGACATTC 259
|||
Db 26 GCTTTTGGGACATTC 41

```

RESULT 24
US-09-258-797-10
: Sequence 10, Application US/09258797
: Patent No. 6183967
: GENERAL INFORMATION:
: APPLICANT: Jayasena, Sumedha
: APPLICANT: Gold, Larry
: TITLE OF INVENTION: Nucleic Acid Ligand I
: FILE REFERENCE: NEX 43C/PCT-CIP
: CURRENT APPLICATION NUMBER: US/09/258,797
: CURRENT FILING DATE: 1999-03-01
: EARLIER APPLICATION NUMBER: 08/945,734

```

; EARLIER FILING DATE: 1997-10-28
; EARLIER APPLICATION NUMBER: 08/487,426
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/487,720
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/484,557
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 78
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-258-797-10

Query Match 1.5%; Score 16; DB 4; Length 78;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 GCTTTGGGACATTC 259
| | | | | | | | | | | | | | | | | |
Db 26 GCTTTGGGACATTC 41

RESULT 25
PCT-US96-09451-10
; Sequence 10, Application PC/TUS9609451
; GENERAL INFORMATION:
; APPLICANT: LARRY GOLD
; APPLICANT: SUMEDHA JAYASENA
; TITLE OF INVENTION: NUCLEIC ACID LIGAND INHIBITORS TO
; TITLE OF INVENTION: DNA POLYMERASES
; NUMBER OF SEQUENCES: 77
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson and Bratschun, L.L.C.
; STREET: 8400 East Prentice Ave., Suite 200
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/09451
; FILING DATE:
; CLASSIFICATION:

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,426
; FILING DATE: 7-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,720
; FILING DATE: 7-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/484,557
; FILING DATE: 7-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: NEX43C/PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 base pairs
; TYPE: nucleic acid

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
PCT-US96-09451-10

Query Match 1.5%; Score 16; DB 5; Length 78;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 GCTTTGGGACATTC 259
| | | | | | | | | | | | | | | | | |
Db 26 GCTTTGGGACATTC 41

RESULT 26
US-09-060-756-531
; Sequence 531, Application US/09060756
; Patent No. 6183957
; GENERAL INFORMATION:
; APPLICANT: Cole, Stewart
; APPLICANT: Buchrieser-Brosch, Roland
; APPLICANT: Gordon, Stephen
; APPLICANT: Billault, Alain
; TITLE OF INVENTION: METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST FROM
; TITLE OF INVENTION: THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED DNA
; FILE REFERENCE: 3495-0169
; CURRENT APPLICATION NUMBER: US/09/060,756
; CURRENT FILING DATE: 1998-04-16
; NUMBER OF SEQ ID NOS: 743
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 531
; LENGTH: 99
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-09-060-756-531

Query Match 1.5%; Score 16; DB 4; Length 99;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 ACCCGGTGGCGGCTG 400
| | | | | | | | | | | | | | | | | |
Db 81 ACCCGGTGGCGGCTG 96

RESULT 27
US-09-060-756-386
; Sequence 386, Application US/09060756
; Patent No. 6183957
; GENERAL INFORMATION:
; APPLICANT: Cole, Stewart
; APPLICANT: Buchrieser-Brosch, Roland
; APPLICANT: Gordon, Stephen
; APPLICANT: Billault, Alain
; TITLE OF INVENTION: METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST FROM
; TITLE OF INVENTION: THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED DNA
; TITLE OF INVENTION: LIBRARY APPLICATION TO THE DETECTION OF MYCOBACTERIA
; FILE REFERENCE: 3495-0169
; CURRENT APPLICATION NUMBER: US/09/060,756
; CURRENT FILING DATE: 1998-04-16
; NUMBER OF SEQ ID NOS: 743
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 386
; LENGTH: 287
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (various positions within the sequence)
; OTHER INFORMATION: applicants are uncertain of bases designated as "n"
US-09-060-756-386

Query Match 1.5%; Score 16; DB 4; Length 287;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 ACCCGGTCGGCGCTG 400
|||||
DB 101 ACCCGGTCGGCGCTG 116

RESULT 28
US-09-060-756-46
; Sequence 46, Application US/09060756
; Patent No. 6183957
; GENERAL INFORMATION:
; APPLICANT: Cole, Stewart
; APPLICANT: Buchrieser-Brosch, Roland
; APPLICANT: Gordon, Stephen
; APPLICANT: Billault, Alain
; TITLE OF INVENTION: METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST FROM
; TITLE OF INVENTION: THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED DNA
; FILE OF INVENTION: LIBRARY APPLICATION TO THE DETECTION OF MYCOBACTERIA
; FILE REFERENCE: 3495-0169
; CURRENT APPLICATION NUMBER: US/09/060,756
; CURRENT FILING DATE: 1998-04-16
; NUMBER OF SEQ ID NOS: 743
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46
; LENGTH: 315
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (various positions within the sequence)
; OTHER INFORMATION: applicants are uncertain of bases designated as "n"
US-09-060-756-46

Query Match 1.5%; Score 16; DB 4; Length 315;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 ACCCGGTCGGCGCTG 400
|||||
DB 88 ACCCGGTCGGCGCTG 103

RESULT 29
US-08-576-202-30/c
; Sequence 30, Application US/08576202
; Patent No. 6350576
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael
; APPLICANT: Lisitsyn, Nikolai
; TITLE OF INVENTION: CANCER DETECTION PROBES
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESS: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: 94111
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/576,202
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015

; REFERENCE/DOCKET NUMBER: A60430-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 335 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "probe"
US-08-576-202-30

Query Match 1.5%; Score 16; DB 4; Length 335;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 GGCAATGGGATCCCCA 1022
|||||
DB 35 GGCAATGGGATCCCCA 20

RESULT 30
PCT-US95-16766-30/c
; Sequence 30, Application PC/TUS9516766
; GENERAL INFORMATION:
; APPLICANT: COLD SPRING HARBOR, Laboratories
; TITLE OF INVENTION: CANCER DETECTION PROBES
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16766
; FILING DATE: 12-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: ROWLAND, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: FP60430-1-PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 335 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "probe"
PCT-US95-16766-30

Query Match 1.5%; Score 16; DB 5; Length 335;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 GGCAATGGGATCCCCA 1022
|||||
DB 35 GGCAATGGGATCCCCA 20

RESULT 31

```
US-08-967-101-122
; Sequence 122, Application US/08967101
; Patent No. 5840540
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HORWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,101
; FILING DATE: 10-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 440 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-967-101-122

Query Match 1.5%; Score 16; DB 2; Length 440;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 ATGATCTGCCTGCCTT 208
|||||
Db 378 ATGATCTGCCTGCCTT 393

RESULT 32
US-08-592-541-122
; Sequence 122, Application US/08592541
; Patent No. 5986054
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HORWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/124,698
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 440 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-592-541-122

Query Match 1.5%; Score 16; DB 2; Length 440;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 ATGATCTGCCTGCCTT 208
|||||
Db 378 ATGATCTGCCTGCCTT 393
```

```
US-08-967-101-122
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,541
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 440 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-592-541-122

Query Match 1.5%; Score 16; DB 2; Length 440;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 ATGATCTGCCTGCCTT 208
|||||
Db 378 ATGATCTGCCTGCCTT 393

RESULT 33
US-09-124-698-122
; Sequence 122, Application US/09124698
; Patent No. 6117978
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HORWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/124,698
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 440 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-124-698-122
```

Query Match 1.5%; Score 16; DB 3; Length 440;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 ATGATCTGCCTGCCTT 208
| | | | | | | | | | | | | | | | | |
Db 378 ATGATCTGCCTGCCTT 393

RESULT 34
US-09-127-480-122
; Sequence 122, Application US/09127480
; Patent No. 6194153
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/127,480
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 440 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-127-480-122

Query Match 1.5%; Score 16; DB 4; Length 440;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 ATGATCTGCCTGCCTT 208
| | | | | | | | | | | | | | | | | |
Db 378 ATGATCTGCCTGCCTT 393

RESULT 35
US-08-496-841C-122
; Sequence 122, Application US/08496841C
; Patent No. 6210919
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 175

CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby, PC
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/496,841C
FILING DATE: 28-Jun-1995
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Paul F. Fehlner, Ph.D.
REGISTRATION NUMBER: 35,135
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 527-7700
TELEFAX: (212) 753-6237

INFORMATION FOR SEQ ID NO: 122:
SEQUENCE CHARACTERISTICS:
LENGTH: 440 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 122:
US-08-496-841C-122

Query Match 1.5%; Score 16; DB 4; Length 440;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 ATGATCTGCCTGCCTT 208
| | | | | | | | | | | | | | | | | |
Db 378 ATGATCTGCCTGCCTT 393

RESULT 36
US-08-967-101-7
; Sequence 7, Application US/08967101
; Patent No. 5840540
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,101
FILING DATE: 10-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/592,541
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 450 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-967-101-7

Query Match 1.5%; Score 16; DB 2; Length 450;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 ATGATCTGCCTGCCTT 208
|||||
Db 383 ATGATCTGCCTGCCTT 398

RESULT 37

US-08-967-101-73
Sequence 73, Application US/08967101
Patent No. 5840540
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,101
FILING DATE: 10-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: 08/592,541
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 73:
SEQUENCE CHARACTERISTICS:
LENGTH: 450 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-967-101-73

Query Match 1.5%; Score 16; DB 2; Length 450;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 ATGATCTGCCTGCCTT 208
|||||
Db 383 ATGATCTGCCTGCCTT 398

RESULT 38

US-08-592-541-7
Sequence 7, Application US/08592541
Patent No. 5986054
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,541
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 450 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-592-541-7

Query Match 1.5%; Score 16; DB 2; Length 450;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 ATGATCTGCCTGCCTT 208
|||||
Db 383 ATGATCTGCCTGCCTT 398

RESULT 39

US-08-592-541-73
Sequence 73, Application US/08592541
Patent No. 5986054
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,541
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 450 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-592-541-73

Query Match 1.5%; Score 16; DB 2; Length 450;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 ATGATCTGCCTGCCTT 208
Db 383 ATGATCTGCCTGCCTT 398

RESULT 40
US-09-124-698-7
; Sequence 7, Application US/09124698
; Patent No. 6117978
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROWMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/124,698
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 450 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-124-698-73
```

Query Match 1.5%; Score 16; DB 3; Length 450;

```
; Best Local Similarity 100.0%; Pred. No. 1.2e+02;
; Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 ATGATCTGCCTGCCTT 208
Db 383 ATGATCTGCCTGCCTT 398

RESULT 41
US-09-124-698-73
; Sequence 73, Application US/09124698
; Patent No. 6117978
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROWMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/124,698
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 450 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-124-698-73

Query Match 1.5%; Score 16; DB 3; Length 450;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 ATGATCTGCCTGCCTT 208
Db 383 ATGATCTGCCTGCCTT 398

RESULT 42
US-09-127-480-7
; Sequence 7, Application US/09127480
; Patent No. 6194153
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROWMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
```

ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/127,480
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/592,541
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 450 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-09-127-480-7

Query Match 1.5%; Score 16; DB 4; Length 450;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 ATGATCTGCCTGCCTT 208
|||||
DB 383 ATGATCTGCCTGCCTT 398

RESULT 43
US-09-127-480-73
Sequence 73, Application US/09127480
Patent No. 6194153
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/127,480
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/592,541
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 73:
SEQUENCE CHARACTERISTICS:
LENGTH: 450 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-127-480-73
Query Match 1.5%; Score 16; DB 4; Length 450;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 ATGATCTGCCTGCCTT 208
|||||
DB 383 ATGATCTGCCTGCCTT 398

RESULT 44
US-08-496-841C-7
Sequence 7, Application US/08496841C
Patent No. 6210919
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 175
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby, PC
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/496,841C
FILING DATE: 28-Jun-1995
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Paul F. Fehlner, Ph.D.
REGISTRATION NUMBER: 35,135
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 527-7700
TELEFAX: (212) 753-6237
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 450 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-08-496-841C-7

Query Match 1.5%; Score 16; DB 4; Length 450;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 ATGATCTGCCTGCCTT 208
|||||
DB 383 ATGATCTGCCTGCCTT 398

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2002, 03:24:48 ; Search time 1727 Seconds

(without alignments)
8135.686 Million cell updates/sec

Title: US-09-886-041-1

Perfect score: 1041

Sequence: 1 atgtacaacgggtgctgctg.....acattgttgagtggaactga 1041

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1041	100.0	1372	11 AF345568	AF345568 Homo sapi
2	32	3.1	306	10 N62053	N62053 EST52H035 W
3	32	3.1	680	12 AG114987	AG114987 Pan trogl
4	32	3.1	852	10 BG677418	BG677418 602625245
5	31	3.0	209	9 AW378847	AW378847 PMO-HT022
6	31	3.0	546	10 BG260219	BG260219 602371328
7	29	2.8	422	9 BB849390	BB849390 BB849390
8	23	2.2	665	12 AG106513	AG106513 Pan trogl
9	20	1.9	119	9 AW876644	AW876644 CM1-PT003
10	20	1.9	396	9 AV734147	AV734147 AV734147
11	20	1.9	507	9 AA659324	AA659324 nu13c07.s
12	20	1.9	509	12 AQ401655	AQ401655 HS_5044_B
13	20	1.9	528	12 AQ637267	AQ637267 RPCI-11-4
14	20	1.9	544	10 BG813280	BG813280 daf73bil.
15	20	1.9	654	9 AA220849	AA220849 mv69405.r
16	20	1.9	879	10 BG823216	BG823216 602726580
17	20	1.9	948	10 BG823067	BG823067 602728095

C 18	20	1.9	1016	12 CMS03CKB	AL237908 Tetraodon
C 19	19	1.8	152	12 AQ426903	AQ426903 CITBI-EI-
C 20	19	1.8	267	9 AV237023	AV237023 AV237023
C 21	19	1.8	268	12 AQ069550	AQ069550 HS_2267_B
C 22	19	1.8	302	10 BE848821	BE848821 uv90b09.y
C 23	19	1.8	305	9 AW528562	AW528562 UI-R-BT1-
C 24	19	1.8	378	9 AI981199	AI981199 pat.pk004
C 25	19	1.8	396	9 AA526211	AA526211 n159c06.s
C 26	19	1.8	399	10 BB478693	BB478693 163050 BA
C 27	19	1.8	417	9 AI510676	AI510676 vx91c07.Y
C 28	19	1.8	437	10 BG995639	BG995639 PM2-HT034
C 29	19	1.8	441	10 BG642257	BG642257 JB5 Droso
C 30	19	1.8	458	12 BH190810	BH190810 TC3-39E7
C 31	19	1.8	463	12 B79552	B79552 CIT-HSP-204
C 32	19	1.8	480	12 AQ682105	AQ682105 HS_2142_B
C 33	19	1.8	481	9 AW928794	AW928794 EST337582
C 34	19	1.8	482	9 AA169173	AA169173 2P20C12.S
C 35	19	1.8	496	9 AA866830	AA866830 vx91c07.f
C 36	19	1.8	508	12 AZ747765	AZ747765 RPCI-24-9
C 37	19	1.8	509	10 BM287008	BM287008 527250 MA
C 38	19	1.8	515	12 AZ253634	AZ253634 RPCI-23-4
C 39	19	1.8	518	9 AA180359	AA180359 zp15g07.s
C 40	19	1.8	518	12 AQ535999	AQ535999 RPCI-11-3
C 41	19	1.8	545	12 AZ239913	AZ239913 RPCI-23-7
C 42	19	1.8	554	9 AW231247	AW231247 JAYL0146.
C 43	19	1.8	556	10 BI677745	BI677745 C1S447 CL
C 44	19	1.8	577	12 AZ041597	AZ041597 RPCI-23-2
C 45	19	1.8	594	12 AQ993291	AQ993291 RPCI-23-3

ALIGNMENTS

RESULT 1
AF345568

LOCUS

DEFINITION

AF345568

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Reference

Authors

Title

Journal

Reference

Authors

Title

Journal

Reference

Authors

Title

Journal

Reference

Authors

Title

Journal

Reference

Authors

Title

Journal

Reference

Authors

Title

Journal

Reference

Authors

Title

Journal

Reference

AF345568 Homo sapiens putative chemokine receptor (FKSG80) mRNA linear HTC 13-JUL-2001 cds.

AF345568.1 GI:13517963

HTC.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1372)

Wang, Y.-g. and Gong, L.

Molecular cloning of FKSG80, a novel gene encoding a putative

chemokine receptor

Unpublished

2 (bases 1 to 1372)

Wang, Y.-g.

Direct Submission

Submitted (06-FEB-2001) Beijing FENGKESHENG Function Gene

Technology Ltd., 4 Tou Tiao Lu Chang Street, Xuanwu District,

Beijing 100050, P.R. China

Location/Qualifiers

1..1372

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="12"

1..1372

/gene="FKSG80"

142..1182

/gene="FKSG80"

/codon_start=1

/product="putative chemokine receptor"

/protein_id="AAK29071.1"

/db_xref="GI:13517964"

/translation="MYNGSCCRIGDTISQVMPPLIVAFVILGALNGVALGCGCFH

KWKSTVFLNLAADFLMICLPRTDYLRRRHWAEGDIPCRVGLFTLNNRAGS

IYFLTVVAADRYFKVYVHPHTVLTALVILGTVLLEHLCLVQVE

TAVSCSFIMESANGWHDIMFQLEFFMPLGIIILFCSEFKIVSLRRRQQLARQARMKKA

TRFMVVAIVITCYLPSVSRARLYFLMTVPSSACDPVSHGALHITLISFTYMNMLDPL
VYFSPSPFKFYKLIKLSLKPQGHKSTQRPSEMPINLGRSRCSISVANFSQS
DQWDHPHIVEWH"

BASE COUNT 289 a 386 c 360 g 337 t
ORIGIN

Query Match 100.0%; Score 1041; DB 11; Length 1372;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTACACAGGGTGTGCTGCGCATCGAGGGGACACCATCTCCAGGTGATGCCGCG 60
Db 142 ATGTACACAGGGTGTGCTGCGCATCGAGGGGACACCATCTCCAGGTGATGCCGCG 201
Qy 61 CTGCTCATGTGGCTTTGCTGCTGGGCGCACTAGGCAATGGGTCGCCCTGTGGTTTC 120
Db 202 CTGCTCATGTGGCTTTGCTGCTGGGCGCACTAGGCAATGGGTCGCCCTGTGGTTTC 261
Qy 121 TGCTTCCACATGAAGACCTGGAAGCCAGCACACTGTTTACCTTTTCAATTTGGCGGTGCT 180
Db 262 TGCTTCCACATGAAGACCTGGAAGCCAGCACACTGTTTACCTTTTCAATTTGGCGGTGCT 321
Qy 181 GATTTCCTCTTATGATGCTGCTGCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
Db 322 GATTTCCTCTTATGATGCTGCTGCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 381
Qy 241 TGGGCTTTTGGGACATTCCTGCTGCGAGTGGGCTTTCACGTTGGCCATGAACAGGCC 300
Db 382 TGGGCTTTTGGGACATTCCTGCTGCGAGTGGGCTTTCACGTTGGCCATGAACAGGCC 441
Qy 301 GGGAGCATCGTGTTCCTTACGCTGGTGGCTCGGACAGGATTTTCAAGTGGTCCACCCC 360
Db 442 GGGAGCATCGTGTTCCTTACGCTGGTGGCTCGGACAGGATTTTCAAGTGGTCCACCCC 501
Qy 361 CACACGGGTGACACTATCTCACCGGGTGGGGCTGGCATCGTGCACCTGTGG 420
Db 502 CACACGGGTGACACTATCTCACCGGGTGGGGCTGGCATCGTGCACCTGTGG 561
Qy 421 GCCCTGGTCATCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCTGTCGAAG 480
Db 562 GCCCTGGTCATCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCTGTCGAAG 621
Qy 481 ACGCCGCTCTCTGTGAGAGCTTCATCATGAGTGGCCCAATGGCTGGCATGACATCATG 540
Db 622 ACGCCGCTCTCTGTGAGAGCTTCATCATGAGTGGCCCAATGGCTGGCATGACATCATG 681
Qy 541 TTCCAGCTGGAGTCTTTATGCCCTCGGCATCATCTTATTTTGTCTCTTCAAGATTGT 600
Db 682 TTCCAGCTGGAGTCTTTATGCCCTCGGCATCATCTTATTTTGTCTCTTCAAGATTGT 741
Qy 601 TGGAGCCTGAGCGGAGGACAGCTGGCCAGACAGGCTCGGATGAAGAGCGCACCCGG 660
Db 742 TGGAGCCTGAGCGGAGGACAGCTGGCCAGACAGGCTCGGATGAAGAGCGCACCCGG 801
Qy 661 TTCAATCATGTGGTGGCAATTTGTTTCATCATGCTACCTACCTGCCAGCGTGTCTGCTAGA 720
Db 802 TTCAATCATGTGGTGGCAATTTGTTTCATCATGCTACCTACCTGCCAGCGTGTCTGCTAGA 861
Qy 721 CTCTATTTCTCTGGACGGTGGCTCGAGTGGCTGCGATCCCTCTGTCATGGGGCCCTG 780
Db 862 CTCTATTTCTCTGGACGGTGGCTCGAGTGGCTGCGATCCCTCTGTCATGGGGCCCTG 921
Qy 781 CACATAACCCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTCTGTTATTATTTT 840
Db 922 CACATAACCCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTCTGTTATTATTTT 981
Qy 841 TCAAGCCCTCTCTTCCCAATTTTACAAAGCTCAAAATCTCAGTCTGAAACCCCAAG 900
Db 982 TCAAGCCCTCTCTTCCCAATTTTACAAAGCTCAAAATCTCAGTCTGAAACCCCAAG 1041
Qy 901 CAGCCAGGACACTCAAAACACAAAGCCGGAAGAGATGCCAATTTGAACTCGGTGCG 960
Db 1042 CAGCCAGGACACTCAAAACACAAAGCCGGAAGAGATGCCAATTTGAACTCGGTGCG 1101

Qy 961 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGTATGGCAATGGATCCC 1020
Db 1102 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGTATGGCAATGGATCCC 1161

Qy 1021 CACATTTGTTAGTGGCACTGA 1041
Db 1162 CACATTTGTTAGTGGCACTGA 1182

RESULT 2
LOCUS NG2053 306 bp mRNA linear EST 12-JUN-1996
DEFINITION EST52h035 WATM1 Homo sapiens cDNA clone 52h035 similar to gb HSHM74
ACCESSION NG2053
VERSION NG2053.1 GI:1209207
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 306)
AUTHORS Bouillaud,F.
TITLE Study of expressed sequences tags in adipose tissue 1995
JOURNAL Unpublished (1995)
COMMENT Contact: Frederic Bouillaud
Centre de Recherche sur l'Endocrinologie moleculaire et le
Developpement
CNRS
9, Rue Jules Hetzel, Meudon Bellevue, 92190 France
Tel: 33 1 45 07 52 87
Fax: 33 1 45 07 58 90
Email: bouillaud@infobiogen.fr
automatic cycle sequencing of PCR amplified insert, single read,
putative coding sequence.
Location/Qualifiers
1..306
/organism="Homo sapiens"
/strain="caucasian"
/db_xref="taxon:9606"
/clone="52h035"
/clone_lib="WATM1"
/note="vector: lambda gtl1; Site_1: EcoRI; White adipose
tissue, sub cutaneous, adult, female. Purification of
polyA mRNA, first strand priming with random
oligonucleotides. Ligation of adaptors EcoRI/NotI, cloning
in EcoRI site of lambda gtl1."
74 a 101 c 68 g 63 t

BASE COUNT 74 a 101 c 68 g 63 t
ORIGIN

Query Match 3.1%; Score 32; DB 10; Length 306;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 790 CTCAGCTTCACCTACATGACAGCATGCTGGA 821
Db 32 CTCAGCTTCACCTACATGACAGCATGCTGGA 63

RESULT 3
LOCUS AG114987/c 680 bp DNA linear GSS 03-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-122B10.F, genomic survey sequence.
ACCESSION AG114987
VERSION AG114987.1 GI:16735506
KEYWORDS GSS; GSS (genome survey sequence).
SOURCE Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male
BAC library clone:PTB-122B10.F.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (sites)

AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE BAC end sequences of Library PTB
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 680)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-cho,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:chumpesgsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI
FEATURES Location/Qualifiers
source
1..680
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="PTB-122B10.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
BASE COUNT 164 a 174 c 200 g 142 t
ORIGIN
Query Match 3.1%; Score 32; DB 12; Length 680;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 790 CTCAGCTTCACCTACATGAACAGCATGCTGGA 821
|||||
Db 230 CTCAGCTTCACCTACATGAACAGCATGCTGGA 199

RESULT 4
BG677418 602625245F1 NCI_CGAP_Skn4 852 bp mRNA linear EST 01-MAY-2001
LOCUS mRNA sequence.
DEFINITION BG677418.1 GI:13908815
ACCESSION BG677418.1
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 852)
/clone_lib="HT0228"
/notes="Organ: head,neck; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 40 a 68 c 44 g 57 t
ORIGIN
Query Match 3.0%; Score 31; DB 9; Length 209;
Best Local Similarity 100.0%; Pred. No. 6.6e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 790 CTCAGCTTCACCTACATGAACAGCATGCTGG 820
|||||
Db 97 CTCAGCTTCACCTACATGAACAGCATGCTGG 127

/clone="IMAGE:4750187"
/clone_lib="NCI_CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: Skin; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: Sall; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 212 a 217 c 229 g 194 t
ORIGIN
Query Match 3.1%; Score 32; DB 10; Length 852;
Best Local Similarity 100.0%; Pred. No. 2.9e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 790 CTCAGCTTCACCTACATGAACAGCATGCTGGA 821
|||||
Db 29 CTCAGCTTCACCTACATGAACAGCATGCTGGA 60

RESULT 5
AW378847 209 bp mRNA linear EST 04-FEB-2000
LOCUS PMO-HT0228-181099-001-d08 HT0228 Homo sapiens cDNA, mRNA sequence.
DEFINITION AW378847
ACCESSION AW378847.1 GI:6883506
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 209)
HCGP http://www.ludwig.org.br/ORESTES.
The FAPESP/LICR Human Cancer Genome Project
Unpublished (1999)
AUTHORS Contact: Simpson A.J.G.
TITLE Laboratory of Cancer Genetics
JOURNAL Ludwig Institute for Cancer Research
COMMENT Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PMO&st2=PMO-HT0228-181099-001-d08&t3=1999-10-18&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 10
High quality sequence stop: 209.
FEATURES Location/Qualifiers
source
1..209
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0228"
/dev_stage="Adult"
/notes="Organ: head,neck; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 40 a 68 c 44 g 57 t
ORIGIN
Query Match 3.0%; Score 31; DB 9; Length 209;
Best Local Similarity 100.0%; Pred. No. 6.6e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 790 CTCAGCTTCACCTACATGAACAGCATGCTGG 820
|||||
Db 97 CTCAGCTTCACCTACATGAACAGCATGCTGG 127

RESULT 6
BG260219
LOCUS
DEFINITION 602371328F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4479391 5', mRNA sequence.
ACCESSION
VERSION BG260219.1 GI:12770035
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 546)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM10311 row: g column: 08
High quality sequence stop: 546.

FEATURES

source
Location/Qualifiers
1..546
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4479391"
/clone_lib="NIH_MGC_93"
/tissue_type="transitional cell papilloma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1-7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
124 a 125 c 149 g 148 t

BASE COUNT

ORIGIN
124 a 125 c 149 g 148 t

Query Match 3.0%; Score 31; DB 10; Length 546;
Best Local Similarity 100.0%; Pred. No. 8.6e+05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 995 GCCAGTCTGATGGCAATGGGATCCGCACAT 1025
|||||
Db 32 GCCAGTCTGATGGCAATGGGATCCGCACAT 62

RESULT 7
BB849390
LOCUS
DEFINITION BB849390 RIKEN full-length enriched, adult inner ear Mus musculus cDNA clone F930010D18 5', mRNA sequence.
ACCESSION
VERSION BB849390.1 GI:17090844
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 422)
AUTHORS Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Imotani, K., Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T., Saito, R., Sakai, K., Sakai, C., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa

TITLE
JOURNAL
COMMENT

Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploring Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:<http://genome.gsc.riken.go.jp/>

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

e mouse tissues.
Location/Qualifiers
1..422
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="F930010D18"
/clone_lib="RIKEN full-length enriched, adult inner ear"
/tissue_type="inner ear"
/dev_stage="adult"
80 a 109 c 119 g 114 t

FEATURES

source
1..422
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="F930010D18"
/clone_lib="RIKEN full-length enriched, adult inner ear"
/tissue_type="inner ear"
/dev_stage="adult"
80 a 109 c 119 g 114 t

Query Match 2.8%; Score 29; DB 9; Length 422;
Best Local Similarity 100.0%; Pred. No. 0.00088;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 337 AGGTATTTCAAAGTGGTCCACCCACCA 365
|||||
Db 146 AGGTATTTCAAAGTGGTCCACCCACCA 174

RESULT 8
AG106513/c
LOCUS
DEFINITION AG106513 Pan troglodytes DNA, clone: PTB-111A21.F, genomic survey sequence.
ACCESSION
VERSION AG106513.1 GI:16727031
KEYWORDS GSS; GSS (genome survey sequence).
SOURCE Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male BAC Library clone:PTB-111A21.F.

ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (sites)
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE BAC end sequences of Library PTB
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 665)

A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Watahiki, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)

AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suehiro-cho,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimps@gs.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
 Vector : pKS145
 R.Site 1 : SacI
 R.Site 2 : SacI
FEATURES source
 1..865 Location/Qualifiers
 /organism="Pan troglodytes"
 /db_xref="taxon:9598"
 /clone="PTB-111A21.F"
 /sex="male"
 /cell_type="lymphoblast"
 /clone_lib="PTB Chimpanzee Male BAC Library"
 181 a 187 c 157 g 137 t 3 others
BASE COUNT
ORIGIN
 Query Match 2.2%; Score 23; DB 12; Length 665;
 Best Local Similarity 100.0%; Pred. No. 1.3;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1002 TGATGGCAATGGGATCCGCACA 1024
 |||||
 Db 660 TGATGGCAATGGGATCCGCACA 638
RESULT 9
 AW876644/c 119 bp mRNA linear EST 22-MAY-2000
LOCUS CM1-PT0030-170100-096-f09 PT0030 Homo sapiens cDNA, mRNA sequence.
DEFINITION AW876644
ACCESSION AW876644
VERSION AW876644.1 GI:8014819
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 119)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovsky-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-cm1-pt0030-170100-096-f09&t3=2000-01-17&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 11

FEATURES source
 High quality sequence stop: 113.
 Location/Qualifiers
 1..119
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="PT0030"
 /dev_stage="Adult"
 /note="Organ: pnet; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 /716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 13 a 46 c 31 g 29 t
BASE COUNT
ORIGIN
 Query Match 1.9%; Score 20; DB 9; Length 119;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 136 ACCTGGAGCCGACACTGT 155
 |||||
 Db 110 ACCTGGAGCCGACACTGT 91
RESULT 10
 AV734147 396 bp mRNA linear EST 17-OCT-2000
LOCUS AV734147 cda Homo sapiens cDNA clone cdaAUC10 5', mRNA sequence.
DEFINITION AV734147
ACCESSION AV734147
VERSION AV734147.1 GI:10851692
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 396)
AUTHORS Yang,Y., Song,H., Peng,Y., Gu,Y., Gao,G., Xiao,H., Xu,X., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu ,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R., Chen,J., Chen,Z. and Han,Z.
TITLE Homo sapiens cDNA cda clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zequang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
 Tel: 86-21-50801919(ex.45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn
 This clone is available at CHGC in Shanghai.
 Location/Qualifiers
 1..396
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="cdaAUC10"
 /clone_lib="cda"
 /tissue_type="pheochromocytoma"
 /dev_stage="Adult"
 /lab_host="BM25.8"
 /note="Vector: pTriplex2; Site_1: sf1A; Site_2: sf1B"
 68 a 83 c 53 g 192 t
BASE COUNT
ORIGIN
 Query Match 1.9%; Score 20; DB 9; Length 396;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 189 CCTTATGATCTGCTGCCTT 208
 |||||
 Db 24 CCTTATGATCTGCTGCCTT 43

```

RESULT 11
AA659324
LOCUS
DEFINITION
    507 bp mRNA linear EST 05-NOV-1997
    n13c07.s1 NCI_CGAP_Pr2 Homo sapiens cDNA clone IMAGE:1207884
    similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION
AA659324
VERSION
AA659324.1 GI:2595478
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 507)
AUTHORS
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
JOURNAL
99380589
MEDLINE
COMMENT
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Research Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 620 row: B column: 23
Seq primer: T7
Class: BAC ends
High quality sequence stop: 509.
FEATURES
source
    Location/Qualifiers
        1..509
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="Plate=620 Col=23 Row=B"
        /clone_lib="RPCI-11 Human Male BAC Library"
        /sex="male"
        /note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
        Male blood DNA was isolated from one randomly chosen donor
        and partially digested with a combination of EcoRI and
        EcoRI Methylase. Size selected DNA was cloned into the
        pBACE3.6 vector at EcoRI sites"
BASE COUNT
106 a 120 c 117 g 163 t
ORIGIN
Query Match 1.9%; Score 20; DB 12; Length 509;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 189 CCTTATGATCGCTGCCTT 208
|||||
Db 271 CCTTATGATCGCTGCCTT 290
|||||

RESULT 13
AA637267
LOCUS
DEFINITION
    528 bp DNA linear GSS 17-JUN-1999
    RPCI-11-479K17.TV RPCI-11 Homo sapiens genomic clone RPCI-11-479K17
    , DNA sequence.
ACCESSION
AQ637267
VERSION
AQ637267.1 GI:5099902
KEYWORDS
GSS.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 528)
AUTHORS
Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
,J.C.
TITLE
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
Unpublished (1997)
JOURNAL
Other_GSSs: RPCI-11-479K17.TJ
COMMENT
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research

AA659324
LOCUS
DEFINITION
    507 bp mRNA linear EST 05-NOV-1997
    n13c07.s1 NCI_CGAP_Pr2 Homo sapiens cDNA clone IMAGE:1207884
    similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION
AA659324
VERSION
AA659324.1 GI:2595478
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 507)
AUTHORS
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
JOURNAL
99380589
MEDLINE
COMMENT
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Research Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 620 row: B column: 23
Seq primer: T7
Class: BAC ends
High quality sequence stop: 509.
FEATURES
source
    Location/Qualifiers
        1..509
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="Plate=620 Col=23 Row=B"
        /clone_lib="RPCI-11 Human Male BAC Library"
        /sex="male"
        /note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
        Male blood DNA was isolated from one randomly chosen donor
        and partially digested with a combination of EcoRI and
        EcoRI Methylase. Size selected DNA was cloned into the
        pBACE3.6 vector at EcoRI sites"
BASE COUNT
106 a 120 c 117 g 163 t
ORIGIN
Query Match 1.9%; Score 20; DB 12; Length 509;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 189 CCTTATGATCGCTGCCTT 208
|||||
Db 271 CCTTATGATCGCTGCCTT 290
|||||

AA637267
LOCUS
DEFINITION
    528 bp DNA linear GSS 17-JUN-1999
    RPCI-11-479K17.TV RPCI-11 Homo sapiens genomic clone RPCI-11-479K17
    , DNA sequence.
ACCESSION
AQ637267
VERSION
AQ637267.1 GI:5099902
KEYWORDS
GSS.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 528)
AUTHORS
Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
,J.C.
TITLE
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
Unpublished (1997)
JOURNAL
Other_GSSs: RPCI-11-479K17.TJ
COMMENT
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research

```


9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeet@ig.org

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genet cs (inforesgen.com). BAC end search page: http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: 77
Class: BAC ends.

FEATURES source

Location/Qualifiers
1..528
/organism="Homo sapiens"
/db_xref="GDB:768380"
/db_xref="taxon:9606"
/clone="RPCI-11-479K17"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPC11 Human Male BAC Library"

BASE COUNT 138 a 156 c 140 g 93 t 1 others

BASE COUNT ORIGIN

Query Match 1.9%; Score 20; DB 12; Length 528;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 898 AAGCAGCCAGGACACTCAA 917
|||||
DB 116 AAGCAGCCAGGACACTCAA 135

RESULT 14 BG813280/c LOCUS

DEFINITION BG813280 544 bp mRNA linear EST 22-MAY-2001
daff73b1l.y1 NICHHD XGC Eye1 Xenopus laevis cDNA clone IMAGE:4755429
5' similar to TR:006744 Q06744 MHC CLASS II HISTOCOMPATIBILITY
ANTIGEN, BETA CHAIN PRECURSOR ;, mRNA sequence.

ACCESSION VERSION KEYWORDS SOURCE ORGANISM

EST.
Xenopus laevis
African clawed frog.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus;

REFERENCE AUTHORS

1 (bases 1 to 544)
Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D.,
Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person,
B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R.,
Waterston, R. and Wilson, R.

TITLE JOURNAL COMMENT

WashU Xenopus EST project, 1999
Unpublished (1999)
Contact: Sandy Clifton, Ph.D.
WashU Xenopus EST project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: Xenopus clones from this library are available
through the I.M.A.G.E. Consortium/LLNL at: infoimage.llnl.gov
Possible reversed clone: similarity on wrong strand
High quality sequence stop: 323.

FEATURES source

Location/Qualifiers
1..544
/organism="Xenopus laevis"

/db_xref="taxon:8355"
/clone="IMAGE:4755429"
/clone_lib="NICHHD_XGC_Eye1"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: This is a Xenopus Gene Collection (XGC
) library."
BASE COUNT 198 a 96 c 92 g 158 t

Query Match 1.9%; Score 20; DB 10; Length 544;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATTGTGGCCTTGTCTGGG 86
|||||
DB 505 ATTGTGGCCTTGTCTGGG 486

RESULT 15 AA220849 LOCUS

DEFINITION AA220849 654 bp mRNA linear EST 10-FEB-1997
mv69d05.r1 Soares mouse 3NME12.5 Mus musculus cDNA clone
IMAGE:660297 5' similar to gb:U5334_rnal CREATINE KINASE, B CHAIN
(HUMAN); gb:M74149 Mouse creatine kinase B gene, complete cds
(MOUSE);, mRNA sequence.

ACCESSION VERSION KEYWORDS SOURCE ORGANISM

AA220849
AA220849.1 GI:1838669
EST.
house mouse.

REFERENCE AUTHORS

1 (bases 1 to 654)
Marras, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

TITLE JOURNAL COMMENT

The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (infoimage.llnl.gov) for further information.
MGI:406145

Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 216.

FEATURES source

Location/Qualifiers
1..654
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:660297"
/clone_lib="Soares mouse 3NME12 5"
/sex="unknown"
/tissue_type="fetus"
/dev_stage="12.5dpc total fetus"
/lab_host="DH10B"

/note="Organ: whole fetus; Vector: pT7T3D-Pac (Pharmacia)
with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5' TGTTACCAATCTGAAGTGGAGCGGCCCTATTTTTTTTTTTTTT
3'], on total mouse RNA [provided by Minoru KO, Wayne
State Univ.]; double-stranded cDNA was ligated to Eco RI

adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 148 a 177 c 193 g 136 t
ORIGIN

Query Match 1.9%; Score 20; DB 9; Length 654;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 507 CTGAGCGCGGAGCGACGCT 626
|||||
Db 10 CTGAGCGCGGAGCGACGCT 29

RESULT 16
BG823216/c
LOCUS 602726580F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4866283 5',
DEFINITION mRNA sequence.
ACCESSION BG823216
VERSION BG823216.1 GI:14170803
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 879)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LUCM1732 row: o column: 20
High quality sequence stop: 797.

FEATURES
source
Location/Qualifiers
1..879

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4866283"
/lab_host="DH10B (phage-resistant)"
/tissue_type="adenocarcinoma cell line"
/note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT 252 a 225 c 193 g 209 t
ORIGIN

Query Match 1.9%; Score 20; DB 10; Length 879;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 189 CCTTATGATCGCTGCCTT 208
|||||
Db 504 CCTTATGATCGCTGCCTT 485

RESULT 17
BG823067/c

LOCUS 602728035F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4867774 5',
DEFINITION mRNA sequence.
ACCESSION BG823067
VERSION BG823067.1 GI:14170654
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 948)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LUCM1736 row: m column: 23
High quality sequence stop: 722.

FEATURES
source
Location/Qualifiers
1..948

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4867774"
/clone_lib="NIH_MGC_15"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT 286 a 233 c 196 g 233 t
ORIGIN

Query Match 1.9%; Score 20; DB 10; Length 948;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 189 CCTTATGATCGCTGCCTT 208
|||||
Db 504 CCTTATGATCGCTGCCTT 485

RESULT 18
CNS03CKB/c

LOCUS CNS03CKB 1016 bp DNA linear GSS 15-MAY-2000
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone 014H16 of library G from Tetraodon nigroviridis, genomic survey sequence.
ACCESSION AL237908
VERSION AL237908.1 GI:7897043
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 1016)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the

```

freshwater pufferfish Tetraodon nigroviridis
Unpublished
2 (bases 1 to 1016)
Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fzames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished
3 (bases 1 to 1016)
Genoscope.
Direct Submission
Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES             source
    1..1016
        /organism="Tetraodon nigroviridis"
        /db_xref="taxon:99883"
        /clone="014H16"
        /clone_lib="G"
        /note="Genoscope sequence ID : C0BG014DD08LPI-end : T7"
BASE COUNT      255 a 281 c 253 g 222 t 5 others
ORIGIN
Query Match          1.9%; Score 20; DB 12; Length 1016;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 606 CCTGAGCGGAGGACGACG 625
      ||||||||||||||||
DB 674 CCTGAGCGGAGGACGACG 655

RESULT 19
AQ426903/c
LOCUS
DEFINITION
AQ426903
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
Map Building
Unpublished (1997)
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.
FEATURES             source
    1..152
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="2572H23"
        /clone_lib="CITBI-E1"
        /sex="male"
freshwater pufferfish Tetraodon nigroviridis
Unpublished
2 (bases 1 to 1016)
Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fzames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished
3 (bases 1 to 1016)
Genoscope.
Direct Submission
Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES             source
    1..1016
        /organism="Tetraodon nigroviridis"
        /db_xref="taxon:99883"
        /clone="014H16"
        /clone_lib="G"
        /note="Genoscope sequence ID : C0BG014DD08LPI-end : T7"
BASE COUNT      255 a 281 c 253 g 222 t 5 others
ORIGIN
Query Match          1.9%; Score 20; DB 12; Length 1016;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 606 CCTGAGCGGAGGACGACG 625
      ||||||||||||||||
DB 674 CCTGAGCGGAGGACGACG 655

RESULT 19
AQ426903/c
LOCUS
DEFINITION
AQ426903
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
Map Building
Unpublished (1997)
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.
FEATURES             source
    1..152
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="2572H23"
        /clone_lib="CITBI-E1"
        /sex="male"

```

```

/cell_type="sperm"
/note="Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
CalTech Human BAC Library D"
BASE COUNT      60 a 22 c 36 g 34 t
ORIGIN
Query Match          1.8%; Score 19; DB 12; Length 152;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 569 GCATCATCTATTTCCTC 587
      ||||||||||||||||
DB 140 GCATCATCTATTTCCTC 122

RESULT 20
AV237023/c
LOCUS
DEFINITION
AV237023
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T.,
Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F.,
Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai
C., Kawai,J., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M.,
Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,
Owa,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibata,K., Shibata
Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H.,
Suzuki,H., Takahashi,F., Tateno,M., Tomimaga,N., Tsunoda,Y.,
Watahiki,A., Watanabe,S., Yamamura,T., Yasunishi,A., Yokota,T.,
Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Konno,H., et al. 1999)
Unpublished (1999)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Sasaki,N., Izawa,M., Watahiki,M., Ozawa,K., Tanaka,T., Yoneda,Y.,
Matsuura,S., Carninci,P., Muramatsu,M., Okazaki,Y. and Hayashizaki
Y.
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
Itoh,M., Kitsuai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
FEATURES             source
    1..267
        /organism="Mus musculus"
        /strain="C57BL/6J"
        /db_xref="taxon:10090"
        /clone="4732417M09"
        /clone_lib="RIKEN full-length enriched, 10 day neonate
        skin"

```


ACCESSION AW528562
 VERSION AW528562.1 GI:7170976
 KEYWORDS EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 305)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 9704477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized corpus-striatum library cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com) The following repetitive elements were found in this cDNA sequence: 25-49,
 >AT_rich#Low_complexity
 Seq primer: M13 Forward
 POLYA=Yes.

FEATURES

source Location/Qualifiers
 1..305
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-BT1-ako-g-08-0-UI"
 /clone_lib="UI-R-BT1"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site1: Not I; Site2: Eco RI; The library UI-R-BT1 is a subtracted library derived from a mixture of the following tissues: hippocampus, thalamus, mid-brain, medulla, corpus striatum, cerebral cortex and testis. For a detailed description of the library from which this clone was derived, please visit our web site at ratstat.eng.uiowa.edu. The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)
 TAG_LIB=UI-R-BT1
 TAG_TISSUE=corpus-striatum
 TAG_SEQ=CTAGG"

BASE COUNT 66 a 59 c 82 g 98 t

Query Match 1.8%; Score 19; DB 9; Length 305;
 Best Local Similarity 100.0%; Pred. NO. 1.3e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 418 TGGGCCCTGGTCATCCTGG 436
 ||||||||||||||||
 Db 167 TGGGCCCTGGTCATCCTGG 185

RESULT 24

AI981199/c
 LOCUS 378 bp mRNA linear EST 07-MAY-2001
 DEFINITION pat.pk0048.b5.f chicken activated T cell cDNA Gallus gallus CDNA clone pat.pk0048.b5.f 5', mRNA sequence.

AI981199
 VERSION AI981199.1 GI:5884227
 KEYWORDS EST.
 SOURCE Chicken.
 ORGANISM Gallus gallus

REFERENCE 1 (bases 1 to 378)
 AUTHORS Tirunaguru,V.G., Sofer,L., Cui,J. and Burnside,J.
 TITLE An expressed sequence tag database of T-cell-enriched activated chicken splenocytes: sequence analysis of 5251 clones

JOURNAL Genomics 66 (2), 144-151 (2000)
 MEDLINE 20318616
 COMMENT Contact: Joan Burnside
 Molecular Endocrinology
 University of Delaware
 40 Townsend Hall, Newark, DE 19717, USA
 Tel: 302 831-1345
 Fax: 302-831-3411
 Email: joan@udel.edu, www.chickest.udel.edu

Seq primer: 7'
 Location/Qualifiers
 1..378
 /organism="Gallus gallus"
 /db_xref="taxon:9031"
 /clone="pat.pk0048.b5.f"
 /clone_lib="chicken activated T cell cDNA"
 /sex="male"
 /cell_type="Con A-activated splenic T cell"
 /lab_host="E.coli TOP10 F'"
 /note="Vector: pCDNA3"

BASE COUNT 116 a 50 c 84 g 108 t 20 others
 ORIGIN

Query Match 1.8%; Score 19; DB 9; Length 378;
 Best Local Similarity 100.0%; Pred. NO. 1.4e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 CTCATTGTGGCCTTTGTGC 82
 ||||||||||||||||
 Db 50 CTCATTGTGGCCTTTGTGC 32

RESULT 25

AI981199/c
 LOCUS 396 bp mRNA linear EST 05-AUG-1997
 DEFINITION ni59c06.sl NCI_CGAP_Ov2 Homo sapiens CDNA clone IMAGE:981130 similar to contains Alu repetitive element; mRNA sequence.

ACCESSION AA526211
 VERSION AA526211.1 GI:2268280
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 396)
 NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 Unpublished (1997)

REFERENCE 1 (bases 1 to 396)
 AUTHORS NCI_CGAP
 TITLE Tumor Gene Index
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: David B. Krizman, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 CDNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/dbp/image/image.html
 Insert Length: 491 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amersham.

FEATURES
source

Location/Qualifiers
1. .396
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:981130"
/clone_lib="NCI_CGAP_Ov2"
/sex="female"
/tissue_type="ovary"
/lab_host="DH10B"
/note="Vector: pAMP10; mRNA made from invasive ovarian tumor, cDNA made by oligo-dT priming. Non-directionally cloned. Size-selected on agarose gel, average insert size 600 bp. Reference: Krizman et al. (1996) Cancer Research 56:5380-5383."

BASE COUNT 76 a 109 c 81 g 130 t

ORIGIN

Query Match 1.8%; Score 19; DB 9; Length 396;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 CCTATGATCGCTGCCT 207
|||||

Db 67 CCTATGATCGCTGCCT 85

RESULT 26

BE478693/C 399 bp mRNA linear EST 28-AUG-2000
LOCUS
DEFINITION 163050 BARC 5BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BE478693
VERSION BE478693.1 GI:9598226
KEYWORDS EST.

SOURCE

ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 399)
AUTHORS Sonstegard,T.S., Capuco,A.V., Van Tassel,C.P., Ashwell,M.S. and Wells,K.D.
TITLE Mapping of Expressed Sequence Tags from a normalized bovine mammary gland cDNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Sonstegard TS
USDA, ARS, Beltsville Agricultural Research Center
Belg-200 Rm 2A, Beltsville, MD 20705, USA
Tel: 301 504 8416
Fax: 301 504 8414
Email: tads@ps1.barc.usda.gov

Single pass sequencing. Bases called and alt-trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCGCCAGTCACGACG
Plate: 8 row: D column: 18
Seq primer: ATTTAGTGACACTATAG.

FEATURES
source

Location/Qualifiers
1. .399
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="BARC 5BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled mRNA isolated from mammary tissues at eight physiological, developmental, and disease states."

BASE COUNT 70 a 119 c 98 g 112 t

ORIGIN

Query Match 1.8%; Score 19; DB 10; Length 399;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 855 TCCCAAAATTCACACAAG 873
|||||

Db 205 TCCCAAAATTCACACAAG 187

RESULT 27

AI510676/c 417 bp mRNA linear EST 12-MAR-1999
LOCUS
DEFINITION vx91c07.y1 Soares_thymus_2NDMT Bos musculus cDNA clone
IMAGE:1282572 5', mRNA sequence.

ACCESSION AI510676
VERSION AI510676.1 GI:4409581
KEYWORDS EST.

SOURCE

ORGANISM house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 417)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:674372

This read is a RESEQUENCE of a previously sequenced mouse clone
correct orientation)
Seq primer: -40RP from Gibco
High quality sequence stop: 402.

FEATURES

Location/Qualifiers
1. .417
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1282572"
/clone_lib="Soares_thymus_2NDMT"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGGTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT773 vector. RNA
provided by Dr. Bertrand Jordan. Library went through two
rounds of normalization, and was constructed by Bento
Soares and M.Fatima Bonaldo."

BASE COUNT 103 a 90 c 111 g 113 t

ORIGIN

Query Match 1.8%; Score 19; DB 9; Length 417;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 422 CCCTGGTCATCCTGGGAAC 440
|||||

Db 287 CCCTGGTCATCCTGGGAAC 269

RESULT 28

BC995639 437 bp mRNA linear EST 13-JUN-2001
LOCUS
DEFINITION PM2-HT0342-190201-008-b04 HT0342 Homo sapiens cDNA, mRNA sequence.
ACCESSION BC995639

VERSION BG995639.1 GI:14399709
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Baia, G. S., Simpson, D. H., Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM2&t2=PM2-HT0342-190201-008-b04&t3=2001-02-19&t4=1>)

Seq primer: puc 18 forward

High quality sequence start: 12

High quality sequence stop: 408.

Location/Qualifiers

1..437

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="HT0342"

/dev_stage="Adult"

/note="Organ: head-neck; Vector: puc18; Site:1: Sma1;
 Site:2: Sma1; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

150 a 71 c 108 g 108 t

BASE COUNT

ORIGIN

Query Match 1.8%; Score 19; DB 10; Length 437;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 CCTGGAAGCCACGACTGT 155

|||||

Db 374 CCTGGAAGCCACGACTGT 392

RESULT 29

BG642257/c

LOCUS

3B5 Drosophila accessory gland pSport1 library Drosophila simulans
 cDNA, mRNA sequence.

ACCESSION BG642257.1 GI:14522113

VERSION EST.

KEYWORDS

SOURCE Drosophila simulans.

ORGANISM

Drosophila simulans

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 441)

Swanson, W.J., Clark, A.G., Waldrip-Dail, H.M., Wolfner, M.F. and

Aquadro, C.F.

TITLE

JOURNAL

MEDLINE

COMMENT

Evolutionary EST analysis identifies rapidly evolving male
 reproductive proteins in *Drosophila*
 Proc. Natl. Acad. Sci. U. S. A. 98 (13), 7375-7379 (2001)
 21309955
 Contact: Swanson WJ
 Molecular Biology and Genetics
 Cornell University
 403 Biotechnology Building, Ithaca, NY 14853-2703, USA
 Tel: 607 254 4839
 Fax: 607 255 6240
 Email: wjs18@cornell.edu.

FEATURES

source

1..441
 Location/Qualifiers
 /organism="Drosophila simulans"
 /db_xref="taxon:7240"
 /clone_lib="Drosophila accessory gland pSport1 library"
 /sex="male"
 /tissue_type="accessory gland"
 /cell_type="reproductive"
 /dev_stage="adult"
 /note="Dissected accessory glands"

BASE COUNT 126 a 92 c 116 g 104 t 3 others

ORIGIN

Query Match 1.8%; Score 19; DB 10; Length 441;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 805 ATGAACAGCATGCTGGATC 823

|||||

Db 141 ATGAACAGCATGCTGGATC 123

RESULT 30

BH190810

LOCUS

DEFINITION

BH190810 458 bp DNA linear GSS 24-OCT-2001

TC3-39E7.TR TC3 Trypanosoma cruzi genomic clone TC3-39E7, DNA

sequence.

ACCESSION BH190810

VERSION BH190810.1 GI:16358086

KEYWORDS GSS.

SOURCE Trypanosoma cruzi.

ORGANISM Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;

Trypanosoma; Schizotrypanum.

REFERENCE 1 (bases 1 to 458)

AUTHORS

Ghedini, E., Malek, J., Shetty, J., Khalak, H., Koo, H., Myler, P., Stuart

, K., Andersson, B. and El-Sayed, N.M.

Use of BAC end sequences from Trypanosoma cruzi CL-Brener TC3

library for gene discovery and map construction

Unpublished (2001)

Other_GSSs: TC3-39E7.TF

Contact: Najib M. El-Sayed

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: nelsayed@tigr.org

Clones are derived from the Trypanosoma cruzi CL-Brener BAC library

TC3. For clone availability, please contact Dr. Bjorn Andersson at

Uppsala University (bjorn.andersson@genpat.uu.se).

Seq primer: M13 Rev

Class: BAC ends.

FEATURES

source

1..458
 Location/Qualifiers

/organism="Trypanosoma cruzi"

/strain="CL Brener"

/db_xref="taxon:5693"

/clone="TC3-39E7"

/clone_lib="TC3"

/note="Vector: pBelBAC11; Site:1: Hin dIII; Constructed

for uppsala University by Marie-Christine Le Paslier in

the laboratory of Denis Le Paslier at the Centre d'Etude du Polymorphisme Humain (CEPH), Paris, France. Briefly, Trypanosoma cruzi CL-Brener agarose embedded DNA (obtained from Dr. Franco da Silveira) was partially digested with Hin dIII. High molecular weight fragments were ligated in pBelO-BAC11 digested with Hin dIII. The average insert size is 100 kb. Total clone coverage: approx. 33 x the haploid genome."

BASE COUNT 102 a 124 c 126 g 106 t
ORIGIN

Query Match 1.88; Score 19; DB 12; Length 458;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 611 GCGGAGGACGACGTGGC 629
|||||

Db 146 GCGGAGGACGACGTGGC 164

RESULT 31

B79552/c

LOCUS 463 bp DNA linear GSS 24-OCT-1998
DEFINITION CIT-HSP-2042N4.TR CIT-HSP Homo sapiens genomic clone 2042N4, DNA sequence.

ACCESSION B79552

VERSION B79552.1 GI:2866575

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 463)

AUTHORS Adams M.D., Rounsley S.D., Zhao S., Field C.E., Bass S., Linher K., Golden K., Berry K., Granger D., Sub E., Wible C., Shizuya H., Simon M. and Venter J.C.

TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map Building (1998)

JOURNAL Unpublished (1998)

COMMENT Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadamsetigr.org

Clones are available from Research Genetics (info@resgen.com). BAC

end search page:

http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html

Seq primer: M13 Reverse

Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..463

/organism="Homo sapiens"

/db_xref="GDB:7053631"

/db_xref="taxon:9606"

/clone="2042N4"

/clone.lib="CIT-HSP"

/sex="Male"

/cell_type="Sperm"

/note="Vector: pBelO-BAC11; Site_1: HindIII; Site_2: HindIII"

BASE COUNT 125 a 121 c 91 g 125 t 1 others

ORIGIN

Query Match 1.88; Score 19; DB 12; Length 463;

Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 668 TGGTGGTGGCAATGTGTT 686
|||||

Db 41 TGGTGGTGGCAATGTGTT 23

RESULT 32

AQ682105/c

LOCUS 480 bp DNA linear GSS 28-JUN-1999

DEFINITION HS-2142_B2_C02_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2142 Col=4 Row=N, DNA sequence.

ACCESSION AQ682105

VERSION AQ682105.1 GI:5258088

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 480)

AUTHORS Mahairas G.G., Wallace J.C., Smith K., Swartzell S., Holzman T., Keller A., Shaker R., Furlong J., Young J., Zhao S., Adams M.D. and Hood L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

MEDLINE 99380589

COMMENT Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones may be purchased from Research Genetics (info@resgen.com).

BAC end Web Server: http://www.htsc.washington.edu

Plate: 2142 row: N column: 4

Seq primer: M13 Reverse

Class: BAC ends

High quality sequence stop: 480.

FEATURES

source

Location/Qualifiers

1..480

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="plate=2142 Col=4 Row=N"

/clone.lib="CIT Approved Human Genomic Sperm Library D"

/sex="Male"

/note="Organ: sperm; Vector: pBelO-BAC11; BAC Clones in E-Coli DH10B"

BASE COUNT 127 a 122 c 100 g 130 t 1 others

ORIGIN

Query Match 1.88; Score 19; DB 12; Length 480;

Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 668 TGGTGGTGGCAATGTGTT 686
|||||

Db 29 TGGTGGTGGCAATGTGTT 11

RESULT 33

AW928794/c

LOCUS 481 bp mRNA linear EST 18-MAY-2001

DEFINITION EST337582 tomato flower buds 8 mm to pre-anthesis, Cornell University Lycopersicon esculentum cDNA clone cTCC3K1 5', mRNA sequence.

ACCESSION AW928794

VERSION AW928794.1 GI:8104189

KEYWORDS EST.

SOURCE tomato.

ORGANISM Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;

Lycopersicon

REFERENCE 1 (bases 1 to 481)

AUTHORS van der Hoeven R.S., Bezzerides J.L., Matern A.L., Holt I.E., Liang F., Upton G., Hansen T., Craven M.B., Bowman C.L., Ahn S., Ronning

TITLE
JOURNAL
COMMENT
C.M., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
Generation of ESTs from tomato flower tissue
Unpublished (1999)

Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.

FEATURES

source

Location/Qualifiers
1. .481
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CTOC3K1"
/clone_lib="tomato flower buds 8 mm to pre-anthesis,
Cornell University"
/tissue_type="flower"
/dev_stage="buds 8mm-to-preanthesis"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Tanksley; Flower buds and flowers were
taken from greenhouse plants (4-8 wks old, TA496). They
were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

BASE COUNT

ORIGIN

155 a 91 c 108 g 127 t
Query Match 1.8%; Score 19; DB 9; Length 481;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 758 ATCCCTCTGTCATGGGCG 776

|||||

Db 303 ATCCCTCTGTCATGGGCG 285

RESULT 34

AA169173

LOCUS

AA169173 482 bp mRNA linear EST 20-DEC-1996
zp20cl2.s1 Stratagene fetal retina 937202 Homo sapiens cDNA clone
IMAGE:610006 3', mRNA sequence.

ACCESSION

AA169173

VERSION

AA169173.1 GI:1747749

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 482)

Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins

,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore

,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,

Schellenberg,K., Soares,M.B., Tan,F., Thierry-Neg,J., Trevisan,E.,

Underwood,K., Wohlmann,P., Waterston,R., Wilson,R. and Marra,M.

Generation and analysis of 280,000 human expressed sequence tags

Genome Res. 6 (9), 807-828 (1996)

9704478

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -40M13 fwd. from Amersham

High quality sequence stop: 404.

Location/Qualifiers

1. .482

/organism="Homo sapiens"

/db_xref="GDB:462526"

/db_xref="taxon:9606"

/clone="IMAGE:610006"
/clone_lib="Stratagene fetal retina 937202"
/sex="mixed"

/lab_host="SOLR (kanamycin resistant)"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:

XhoI; Cloned unidirectionally. Primer: Oligo dT. Pooled

retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR

Vector; -5' adaptor sequence: 5' GAATTCGGCAG 3' -3'

adaptor sequence: 5' CTCAGTTTTTTTTTTTTTTT 3'

BASE COUNT 122 a 146 c 73 g 141 t

ORIGIN

Query Match 1.8%; Score 19; DB 9; Length 482;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 851 CCTTCCCAAAATCTACAA 869

|||||

Db 286 CCTTCCCAAAATCTACAA 304

RESULT 35

AA866830/c

LOCUS

AA866830 496 bp mRNA linear EST 16-MAR-1998
vx91c07.r1 Soares_thymus_2NbMT Mus musculus cDNA clone
IMAGE:1282572 5', mRNA sequence.

ACCESSION

AA866830

VERSION

AA866830.1 GI:2962275

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Mus musculus

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 496)

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,

Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,

Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,

Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and

Waterston,R.

The WashU-HMI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:674372

Seq primer: -28m13 rev2 ET from Amersham

High quality sequence stop: 491.

Location/Qualifiers

1. .496

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="IMAGE:1282572"

/clone_lib="Soares_thymus_2NbMT"

/sex="male"

/tissue_type="Thymus"

/dev_stage="4 weeks"

/lab_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5'

TGTTACCAATCTGAAGTGGGAGCGCGGTTTTTTTTTTTTTTTTTTT

3']; double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT73 vector. RNA

provided by Dr. Bertrand Jordan. Library went through two

rounds of normalization, and was constructed by Bento

```

BASE COUNT      117 a   106 c   136 g   137 t
ORIGIN
Query Match      1.8%; Score 19; DB 9; Length 496;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 422 CCTGTGTCATCCTGGGAAC 440
|||||
Db 287 CCTGTGTCATCCTGGGAAC 269

RESULT 36
AZ747765
LOCUS      508 bp   DNA   linear   GSS 25-JAN-2001
DEFINITION RPCI-24-99M13.TJ RPCI-24 Mus musculus genomic clone RPCI-24-99M13,
DNA sequence.
ACCESSION  AZ747765
VERSION     AZ747765.1 GI:12531931
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 508)
AUTHORS    Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
Russell,D., de Jong,P. and Fraser,C.M.
TITLE      Mouse BAC End Sequences from Library RPCI-24
JOURNAL    Unpublished (1999)
COMMENT    Other-GSSs: RPCI-24-99M13.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC
Library availability, please contact Pieter de Jong
{pdejong@email.cho.org}. Clones may be purchased from BACPAC
Resources (http://www.choi.org/bacpac/orderingframe.htm). BAC end
page: http://www.tigr.org/tdb/bac\_ends/mouse/bac\_end\_intro.html
Plate: 99 row: M column: 13
Seq primer: SP6
Class: BAC ends.
FEATURES             Location/Qualifiers
     source           1..508
     strain="C57BL/6J"
     db_xref="taxon:10090"
     clone_lib="RPCI-24-99M13"
     /clone_lib="RPCI-24"
     /sex="Male"
     /cell_type="Spleen/Brain"
     /note="Vector: pTARBAC1; Site.1: BamHI; Site.2: BamHI;
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pTARBAC1 cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."
BASE COUNT      168 a   89 c   97 g   154 t
ORIGIN
Query Match      1.8%; Score 19; DB 12; Length 508;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 907 GGACACTCAAAAACACAAA 925
|||||
Db 232 GGACACTCAAAAACACAAA 250

Soares and M.Fatima Bonaldo."
BASE COUNT      117 a   106 c   136 g   137 t
ORIGIN
Query Match      1.8%; Score 19; DB 9; Length 496;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 422 CCTGTGTCATCCTGGGAAC 440
|||||
Db 287 CCTGTGTCATCCTGGGAAC 269

RESULT 37
BM287008/c
LOCUS      509 bp   mRNA   linear   EST 28-DEC-2001
DEFINITION 527250 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION  BM287008
VERSION     BM287008.1 GI:17996034
KEYWORDS   EST.
SOURCE     COW.
ORGANISM   Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 509)
AUTHORS    Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
Pertea,G., Holt,I., Karanycheva,S., Liang,F., Quackenbush,J. and
Keele,J.W.
TITLE      Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
JOURNAL    Genome Res. 11 (4), 626-630 (2001)
COMMENT    21180013
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCAGTCACGACG
Plate: 135 row: J column: 23
Seq primer: ATTAGGTGACACTATAG.
Location/Qualifiers
     source           1..509
     /organism="Bos taurus"
     /db_xref="taxon:9913"
     /clone_lib="MARC 3BOV"
     /tissue_type="pooled"
     /lab_host="DH10B"
     /note="Vector: pCMV SPORT6; Site.1: XbaI; Site.2: XhoI;
Library made from pooled tissue from marrow, alveolar
macrophage, ovary, fetal semitendinosus muscle, and fetal
longissimus muscle."
BASE COUNT      115 a   144 c   133 g   117 t
ORIGIN
Query Match      1.8%; Score 19; DB 10; Length 509;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 855 TCCCAAATTCACAACAAG 873
|||||
Db 430 TCCCAAATTCACAACAAG 412

RESULT 38
AZ253634/c
LOCUS      515 bp   DNA   linear   GSS 15-JUN-2000
DEFINITION RPCI-23-42F8.TV RPCI-23 Mus musculus genomic clone RPCI-23-42F8,
DNA sequence.
ACCESSION  AZ253634
VERSION     AZ253634.1 GI:8566837
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 515)
AUTHORS    Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret

```

TITLE
 JOURNAL
 COMMENT
 B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P. and Fraser, C.M.
 Mouse BAC End Sequences from Library RPCI-23
 Unpublished (1999)
 Other_GSSs: RPCI-23-42F8.TJ
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html
 Plate: 42 row: F column: 8
 Seq primer: T7
 Class: BAC ends.

FEATURES

source

Location/Qualifiers

1. .515
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-23-42F8"
 /clone_lib="RPCI-23"
 /sex="Female"
 /lab_host="DH10B"
 /note="Organ: Kidney/Brain; Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACE3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."
 151 a 126 c 100 g 138 t

BASE COUNT
 ORIGIN

Query Match 1.8%; Score 19; DB 12; Length 515;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 699 CTTGCCAGCGTCTGCT 717
 Db 273 CTTGCCAGCGTCTGCT 255

RESULT 39

AA180359

LOCUS

DEFINITION 518 bp mRNA linear EST 31-DEC-1996
 zplsg07.s1 Stratagene fetal retina 937202 Homo sapiens cDNA clone
 IMAGE:609564 3', mRNA sequence.

ACCESSION

AA180359

VERSION

AA180359.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 518)

Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,

Chisoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W., Hawkins

, M., Hultman, M., Kucaba, F., Lacy, M., Le, M., Le, N., Marais, E., Moore

, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,

Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E.,

Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.

Generation and analysis of 280,000 human expressed sequence tags

Genome Res. 6 (9), 807-828 (1996)

97044478

TITLE

JOURNAL

MEDLINE

COMMENT

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: estewatson.wustl.edu

This clone is available royalty-free through LLMU; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -40M13 fwd. from Amersham

High quality sequence stop: 382.

Location/Qualifiers

1. .518

/organism="Homo sapiens"

/db_xref="GDB:4624823"

/db_xref="taxon:9606"

/clone="IMAGE:609564"

/clone_lib="Stratagene fetal retina 937202"

/sex="mixed"

/lab_host="SOLR (kanamycin resistant)"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:

XhoI; Cloned unidirectionally. Primer: Oligo dT. Pooled

retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR

Vector: -5' adaptor sequence: 5' GAATTCGGCAGG 3' -3'

adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'

131 a 150 c 84 g 151 t 2 others

BASE COUNT

ORIGIN

Query Match 1.8%; Score 19; DB 9; Length 518;

Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 851 CCTTCCCAATTCACAA 869

Db 286 CCTTCCCAATTCACAA 304

RESULT 40

AAQ535999/c

LOCUS

DEFINITION 518 bp DNA linear GSS 18-MAY-1999

RPCI-11-318M7.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-318M7,

DNA sequence.

ACCESSION AQ535999

VERSION AQ535999.1

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 518)

Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter

, J.C.

Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready

Map Building

Unpublished (1997)

Other_GSSs: RPCI-11-318M7.TV

Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: hbe@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or fromResearch Genet cs (info@resgen.com). BAC end search page:http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html.

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

1. .518

/organism="Homo sapiens"

/db_xref="GDB:7622022"

/db_xref="taxon:9606"

FEATURES

source

```

/clone="RPCI-11-318M7"
/clone_lib="RPCI-11"
/sex="Male"
/notes="Lymphocytes"
/notes="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"
BASE COUNT      168 a      81 c      92 g      177 t
ORIGIN

Query Match
Best Local Similarity 1.8%; Score 19; DB 12; Length 518;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 189 CCTTATGATCGCTGCT 207
|||||
Db 187 CCTTATGATCGCTGCT 169

RESULT 41
A2239913/c
LOCUS      A2239913      545 bp      DNA      linear      GSS 15-JUN-2000
DEFINITION RPCI-23-73N16.TJ RPCI-23 Mus musculus genomic clone RPCI-23-73N16,
DNA sequence.
ACCESSION  A2239913
VERSION    A2239913.1 GI:8553104
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE  1 (bases 1 to 545)
AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
and Fraser,C.M.
TITLE     Mouse BAC End Sequences from Library RPCI-23
JOURNAL   Unpublished (1999)
COMMENT   Other_GSSs: RPCI-23-73N16.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buhalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buhalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 73 row: N column: 16
Seq primer: SP6
Class: BAC ends.
FEATURES             location/Qualifiers
source               1..545
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-73N16"
/clone_lib="RPCI-23"
/sex="Female"
/lab_host="DH10B"
/notes="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methyase. Size
selected DNA was cloned into the pBACe3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT      154 a      135 c      112 g      144 t
ORIGIN

Query Match
Best Local Similarity 1.8%; Score 19; DB 12; Length 545;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 566 TCGGCATCATCTTATTG 584
|||||
Db 375 TCGGCATCATCTTATTG 393

RESULT 43
BI677745/c
LOCUS      BI677745      556 bp      mRNA      linear      EST 17-SEP-2001
DEFINITION CUS447 CLS (Cambium and bark region of black locust - Summer)

```

```

Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 699 CTTGCCAGCGTGTCTGCT 717
|||||
Db 308 CTTGCCAGCGTGTCTGCT 290

RESULT 42
AW231247
LOCUS      AW231247      554 bp      mRNA      linear      EST 10-DEC-1999
DEFINITION JAYL0146.GYL Schistosoma japonicum Lambda gtl1 Express library
Schistosoma japonicum cDNA clone JAYL0146.GY 5', mRNA sequence.
ACCESSION  AW231247
VERSION    AW231247.1 GI:6560543
KEYWORDS   EST.
SOURCE     Schistosoma japonicum.
ORGANISM   Schistosoma japonicum
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
Strigoidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
REFERENCE  1 (bases 1 to 554)
AUTHORS   Li,Y., Wu,Z.D. and Yu,X.B.
TITLE     Expressed sequence tags from adults of Schistosoma japonicum
          (Chinese strain) (Li,Y.; Wu,Z.D.; Yu,X.B.)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Wu ZD
          Department of Parasitology
          Sun-Yat-sen University of Medical Sciences
          BOX 510089, 74# Zhongshen Er Road, Guangzhou, Guangdong, P.R.China
          Tel: 86-20-87330566
          Fax: 86-20- 87331679
          Email: zdwu62@163.net
          PCR Primers
          FORWARD: Lambda gtl1 Forward Primer
          BACKWARD: Lambda gtl1 Reverse Primer
          Seq primer: Lambda gtl1 Forward Primer
          High quality sequence stop: 554.
FEATURES             Location/Qualifiers
source               1..554
/organism="Schistosoma japonicum"
/strain="Chinese"
/db_xref="taxon:6182"
/clone="JAYL0146.GY"
/clone_lib="Schistosoma japonicum Lambda gtl1 Express
library"
/sex="Mix"
/notes="Vector: Lambda gtl1 Sfi-Not; Site_1: EcoRI; Site_2:
NotI; Several hundred adult Schistosoma japonicum(Jiangxi,
P.R.China, strain), of mixed sex, were perfused from the
mesenteries of experimentally infected rabbits.
Double-strain cDNA synthesized with the mRNA isolated
from adult worm, was inserted into the bacteriophage
lambda gtl1 Sfi-Not arms between EcoRI and NotI site of
the LacZ gene. The cDNA library was constructed by Chen
S.Z. at Nanjing Medical University, Nanjing, Jiangsu,
P.R. China. (see: Chen Shuzhen, et al. Chinese Journal of
Zoonoses 1997,13(6): 23-25)"
BASE COUNT      203 a      115 c      79 g      157 t
ORIGIN

Query Match
Best Local Similarity 1.8%; Score 19; DB 9; Length 554;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 566 TCGGCATCATCTTATTG 584
|||||
Db 375 TCGGCATCATCTTATTG 393

RESULT 43
BI677745/c
LOCUS      BI677745      556 bp      mRNA      linear      EST 17-SEP-2001
DEFINITION CUS447 CLS (Cambium and bark region of black locust - Summer)

```

Robinia pseudoacacia cDNA, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BI677745
GI:15630652
EST.
Robinia pseudoacacia.
Robinia pseudoacacia

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Robinieae;
Robinia.

REFERENCE
AUTHORS

1 (bases 1 to 556)
Han,K.-H., Yang,J., Park,S., Paule,C.R., Kapur,V., Retzel,E.F.,
Kamdem,D.P. and Keathley,D.E.

TITLE

Analysis of gene expression patterns in trunk wood of a mature

JOURNAL

black locust (Robinia pseudoacacia)

COMMENT

Unpublished (2002)

Contact: Kyung-Hwan Han

Department of Forestry

Michigan State University

126 Natural Resources, East Lansing, MI 48824-1222, USA

Tel: 517 353 4751

Fax: 517 432 1143

Email: hanky@msu.edu.

FEATURES

Location/Qualifiers

source

1..556

/organism="Robinia pseudoacacia"

/db_xref="taxon:35938"

/clone_lib="CLS (Cambium and bark region of black locust -

Summer)"

/tissue_type="cambial region (Cambium and bark region of

black locust - Summer)"

/dev_stage="mature tree"

/note="Vector: lambda Triplex; Site_1: Sfi IA; Site_2: Sfi

IB; The cDNA library was made from the cambium and bark

region of a mature black locust tree collected in Michigan

in late July."

148 a 99 c 134 g 175 t

BASE COUNT

ORIGIN

Query Match 1.8%; Score 19; DB 10; Length 556;

Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 GCTTTGGGACATTCCT 262

|||||

Db 169 GCTTTGGGACATTCCT 151

RESULT 44

AZ041597/c

LOCUS

DEFINITION

RPIC-23-286K10.TJ RPIC-23 Mus musculus genomic clone RPIC-23-286K10

, DNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 577)

Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret

,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.

and Fraser,C.M.

Mouse BAC End Sequences from Library RPIC-23

Unpublished (1999)

Other_GSSs: RPIC-23-286K10.TV

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

*Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPIC-23. For BAC
library availability, please contact Pieter de Jong
(pieter@jlong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 286 row: K column: 10
Seq primer: SP6
Class: BAC ends.

FEATURES

Location/Qualifiers

source

1..577

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="RPIC-23-286K10"

/clone_lib="RPIC-23"

/sex="Female"

/lab_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBACe3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 190 a 122 c 95 g 170 t

ORIGIN

Query Match 1.8%; Score 19; DB 12; Length 577;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 811 AGCATGCTGGATCCCTGG 829

|||||

Db 263 AGCATGCTGGATCCCTGG 245

RESULT 45

AQ993291/c

LOCUS

DEFINITION

RPIC-23-331E8.TVB RPIC-23 Mus musculus genomic clone RPIC-23-331E8,

DNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 594)

Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret

,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.

and Fraser,C.M.

Mouse BAC End Sequences from Library RPIC-23

Unpublished (1999)

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPIC-23. For BAC

library availability, please contact Pieter de Jong

(pieter@jlong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)

or from Resea ch Genetics (info@resgen.com). BAC end page:

http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html

Plate: 331 row: E column: 8

Seq primer: T7

Class: BAC ends.

FEATURES

Location/Qualifiers

```
source      1. .594
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-331E8"
/clone_lib="RPCI-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBACe3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT 180 a 145 c 113 g 154 t 2 others
ORIGIN
Query Match      1.8%; Score 19; DB 12; Length 594;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 699 CCTGCCCCAGCGTGTCTGCT 717
      |||||
Db 295 CCTGCCCCAGCGTGTCTGCT 277
```

Search completed: October 29, 2002, 04:51:15
Job time : 1756 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2002, 02:40:58 ; Search time 1723 Seconds
(without alignments)
8154.574 Million cell updates/sec

Title: US-09-886-041-1
Perfect score: 1041
Sequence: 1 atgtacaacgggtcgtgctg.....acattgttgatggcactga 1041

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estcov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1041	100.0	1372	11	AF345568	Homo sapi
2	317.2	30.5	870	9	AL545172	AL545172
3	317.2	30.5	885	9	AL546894	AL546894
4	309	29.7	801	10	BI837965	BI837965
5	289.4	27.8	855	9	AL554198	AL554198
6	278.6	26.0	876	10	BI820995	BI820995
7	270.4	26.0	734	10	BI908162	BI908162
8	267.2	25.7	752	10	BM008116	BM008116
9	255.8	24.6	367	9	BB873518	BB873518
10	254.6	24.5	660	9	BB614132	BB614132
11	253.2	24.3	367	9	BB873597	BB873597
12	230	22.1	422	9	BB849390	BB849390
13	218.2	21.0	550	10	BF074336	BF074336
14	215.8	20.7	636	10	BI489649	BI489649
15	213.8	20.5	532	10	BF230383	BF230383
16	208.6	20.0	938	10	BI521556	BI521556
17	196.6	18.9	509	9	BB866696	BB866696

18	195.4	18.8	689	9	BB613494	BB613494	
19	194.6	18.7	479	10	BI961009	BI961009	
20	183	17.6	819	10	BI822957	BI822957	
21	176.8	17.0	638	9	BB661057	BB661057	
22	169.4	16.3	647	9	BB645691	BB645691	
23	166.8	16.0	1016	12	CNS05PYK	AL348581 Tetraodon	
C	24	162.4	15.6	680	12	AG114987	AG114987 Pan trogl
	25	137.2	13.2	405	9	AW801664	AW801664 IL5-UM006
C	26	123.8	11.9	649	12	A2287727	A2287727 RPCI-23-4
	27	122.6	11.8	1017	10	BI906118	BI906118 603062522
28	122.2	11.7	658	9	BB636095	BB636095	
29	118	11.3	406	10	BE718299	BE718299 CM3-HT080	
30	105.6	10.1	805	10	BG742555	BG742555 602632616	
C	31	105	10.1	476	12	A2882355	A2882355 RPCI-23-2
32	102.6	9.9	894	12	CNS02RR8	AL210941 Tetraodon	
33	100.2	9.6	2020	11	EC013202	EC013202 Homo sapi	
34	97	9.3	932	12	CNS025IR	AL211932 Tetraodon	
35	92	8.8	663	10	BM426517	BM426517 pgf2n.pK0	
36	90.4	8.7	696	10	BJ011416	BJ011416	
37	87.6	8.4	734	10	BG914321	BG914321 602810633	
38	85.8	8.2	605	10	BI960974	BI960974 MON01_3_D	
39	85.8	8.2	620	10	BI961697	BI961697 MON01_3_D	
40	85.2	8.2	609	9	AL588350	AL588350	
41	84.2	8.1	618	10	BF346620	BF346620 602021219	
42	84.2	8.1	663	10	BG820137	BG820137 602782353	
43	84.2	8.1	691	10	BG912683	BG912683 602807060	
44	84.2	8.1	703	10	BF527616	BF527616 602040566	
45	84.2	8.1	764	10	BF345996	BF345996 602018717	

ALIGNMENTS

RESULT 1
AF345568
LOCUS Homo sapiens putative chemokine receptor (FKSG80) mRNA, complete cds.
DEFINITION AF345568.1 GI:13517963
ACCESSION AF345568
VERSION AF345568.1
KEYWORDS HTC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1372)
AUTHORS Wang,Y.-g. and Gong,L.
TITLE Molecular cloning of FKSG80, a novel gene encoding a putative chemokine receptor
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1372)
AUTHORS Wang,Y.-g.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-2001) Beijing FENGKESHENG Function Gene Technology Ltd., 4 Tou Tiao Lu Chang Street, Xuanwu District, Beijing 100050, P.R. China
FEATURES
source Location/Qualifiers
1..1372
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
1..1372
/gene="FKSG80"
142..1182
/gene="FKSG80"
/product="putative chemokine receptor"
/protein_id="AAK29071.1"
/db_xref="GI:13517964"
translation="MTNGSCCRIGEDTISQVMPPLLIIVAFVLGALNGVALCGCFPHM
KWKSTVYLFNLAVADFLMICLPFTDYLRRRHWAFGDIICRVGFTLANNRAGS
IYFLVAVRADRYKVVHPHHVITSTRAAGIVCTLWALVILGVLLEHLNVQVE
TAVSCSFIMESANGWHDIMFQLEFFMPLGIILFCSEFKIVSLRRRQQLARQARMKA

TRFMVVAIVITCYLPSVSARLYFLMTVPSSACDPSVHGALHITLSFTYWNMLDPL
VYFSSPFFKFKYKIKSLKPKQPHGSKTQREEMPISNLGRRSCISVANSFOSQS
DQOWDPHIVSWH*

BASE COUNT 289 a 386 c 360 g 337 t
ORIGIN

Query Match 100.0%; Score 1041; DB 11; Length 1372;
Best Local Similarity 100.0%; Pred. No. 1.1e-274;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTACACGGGTGCTGCTGGCATCGAGGGGACACCATCTCCAGGTATGCCCGC 60
Db 142 ATGTACACGGGTGCTGCTGGCATCGAGGGGACACCATCTCCAGGTATGCCCGC 201
Qy 61 CTGCTCATGTGGCTTTGTCTGGCGCACTAGCAATGGGTCGCCCTGTGGTTTC 120
Db 202 CTGCTCATGTGGCTTTGTCTGGCGCACTAGCAATGGGTCGCCCTGTGGTTTC 261
Qy 121 TGCCTCCACATGAAGACCTGGAAGCCAGCACACTGTTTACCTTTTCAATTTGGCCGTGGCT 180
Db 262 TGCCTCCACATGAAGACCTGGAAGCCAGCACACTGTTTACCTTTTCAATTTGGCCGTGGCT 321
Qy 181 GATTTCCTCTTATGATCTGCTGCGGACACTATTACCTCAGACCTAGACAC 240
Db 322 GATTTCCTCTTATGATCTGCTGCGGACACTATTACCTCAGACCTAGACAC 381
Qy 241 TGGGCTTTTGGGACATTCCTCGCGAGTGGGCTCTTACGTTGGCCATGAACAGGCC 300
Db 382 TGGGCTTTTGGGACATTCCTCGCGAGTGGGCTCTTACGTTGGCCATGAACAGGCC 441
Qy 301 GGGAGACGCTGTTCTTACGGTGGTGCGTGGGACAGGATTTTCAAAGTGGTCCACCCC 360
Db 442 GGGAGACGCTGTTCTTACGGTGGTGCGGACAGGATTTTCAAAGTGGTCCACCCC 501
Qy 361 CACCACGGGTGAACACTATCTCCACCGGCTGGCGGTGCATCGTGCACCTGTGG 420
Db 502 CACCACGGGTGAACACTATCTCCACCGGCTGGCGGTGCATCGTGCACCTGTGG 561
Qy 421 GCCCTGGTCATCTCTGGGAACAGTGTATCTTTGTGGAGAACCATCTCTGGTGCACAG 480
Db 562 GCCCTGGTCATCTCTGGGAACAGTGTATCTTTGTGGAGAACCATCTCTGGTGCACAG 621
Qy 481 ACGGCGGTCTCCTGTGAGAGCTTCATCATGAGTGGGCCAATGGCTGGCATGACATCATG 540
Db 622 ACGGCGGTCTCCTGTGAGAGCTTCATCATGAGTGGGCCAATGGCTGGCATGACATCATG 681
Qy 541 TTCCAGCTGGAGTCTTTATGCCCTCGGCATCATCTTATTTTGTCTTCAAGATTGTT 600
Db 682 TTCCAGCTGGAGTCTTTATGCCCTCGGCATCATCTTATTTTGTCTTCAAGATTGTT 741
Qy 601 TGGAGCCTGAGCGGAGGACAGCTGGCCAGACAGGCTCGGATGAAGAGCGGACCCGG 660
Db 742 TGGAGCCTGAGCGGAGGACAGCTGGCCAGACAGGCTCGGATGAAGAGCGGACCCGG 801
Qy 661 TTCAATCATGTGGTGGCAATTTGTTCATCATCATGCTACCTGCCAGCGTGTCTGTAGA 720
Db 802 TTCAATCATGTGGTGGCAATTTGTTCATCATCATGCTACCTGCCAGCGTGTCTGTAGA 861
Qy 721 CTCTATTTCCTCTGGACGGTGGCTCGAGTGGCTGCGATCCCTCTGTCCATGGGGCCCTG 780
Db 862 CTCTATTTCCTCTGGACGGTGGCTCGAGTGGCTGCGATCCCTCTGTCCATGGGGCCCTG 921
Qy 781 CACATAACCCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTCTGTATATTATTT 840
Db 922 CACATAACCCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTCTGTATATTATTT 981
Qy 841 TCAAGCCCTCTCTTCCCAATTTCTACAACAAGCTCAAAATCTCAGTCTGAAACCCCAAG 900
Db 982 TCAAGCCCTCTCTTCCCAATTTCTACAACAAGCTCAAAATCTCAGTCTGAAACCCCAAG 1041
Qy 901 CAGCCAGACACTCAAAACACAAAGCCGGAAGAGATGCCAATTTTGAACCTCGGTGCG 960
Db 1042 CAGCCAGACACTCAAAACACAAAGCCGGAAGAGATGCCAATTTTGAACCTCGGTGCG 1101

Qy 961 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGTATGGCAATGGATCCC 1020
Db 1102 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGTATGGCAATGGATCCC 1161

Qy 1021 CACATTGTTGAGTGGCACTGA 1041
Db 1162 CACATTGTTGAGTGGCACTGA 1182

RESULT 2
AL545172 AL545172 LTI_NFL006_PL2 870 bp mRNA linear EST 16-FEB-2001
LOCUS
DEFINITION
prime, mRNA sequence.

ACCESSION
AL545172
VERSION
AL545172.1
KEYWORDS
EST.
SOURCE
human.

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 870)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
CONTACT
Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..870
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI028YL20"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 169 a 244 c 229 g 227 t 1 others
ORIGIN

Query Match 30.5%; Score 317.2; DB 9; Length 870;
Best Local Similarity 65.0%; Pred. No. 4.7e-76;
Matches 485; Conservative 0; Mismatches 258; Indels 3; Gaps 1;

Qy 1 ATGTACACGGGTGCTGCTGGCGCACTAGGCGGACACCATCTCCAGGTATGCCCGCG 60
Db 89 ATACACAAGAAGAACTGCTGTGCTTCCGAGATGACTTTCATTGTCAAGGTCTTCCCGCG 148

Qy 61 CTGCTCATTTGGGCTTTGTGCTGGCGCACTAGGCGGACACCATCTCCAGGTATGCCCGCG 120
Db 149 GTGTTGGGGCTGGAGTTTATCTTCGGGCTTCTGGGCAATGGGCTTGGCCCTGTGATTTC 208

Qy 121 TGCCTCCACATGAAGAACTTGAAGCCAGCACTGTTTACCTTTTCAATTTGGCCGTGGCT 180
Db 209 TGTTCACCTCAAGTCTTGGAAATCCAGCGGATTTCTGTTCAACCTGGCAGTGGCT 268

Qy 181 GATTTCCTCTTATGATCTGCCCTTTTGGAGACAGTATTTACCTCAGACGTAGACAC 240
Db 269 GACTTTCTACTGATCATCTGCTGCCCTTCTGATGGACAACTATGTAGGCGGTGGGAC 328

Qy 241 TGGGCTTTGGGACATTCCTGCCGAGTGGGCTCTTCAGCTTGGCCATGAACAGGCC 300
Db 329 TGGAGTTTGGGACATTCCTTGGCGGTGATGCTCTTCAATTTGGCTATGAACCGCCAG 388


```
QY 301 GGGAGCATCGTGTCTACGGTGGTGGCTGGGACAGAGTATTTCACAACTGGTCCACCCC 360
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 389 GGCAGCATCATCTCTCAGGTGGTGGGGTAGACAGAGTATTTCGGGTGGTCCATCCC 448
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 CACCACGGGTGAACACTATCTCCACCGGGTGGGGCTGGCATCGTCTGCACCCCTGTGG 420
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 449 CACCACGGCTGAACAAGATCTCCAAATCGACAGCAGCCATCATCTCTTGGCTTCGTGG 508
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 GCCCTGGTCATCTCTGGACAGTGTATCTTTTCTGGGAAACCATCTCTGGCTGCAAGAG 480
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 509 GGCATCACTATTGGCTGACAGTCCACCTCCTGAAGAAGAAGATGCCGATCCAGATGCG 568
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 481 ACGCGCTCTCTCTGTCAGAGCTTCATCATGGAGTGGCGCATGGCTGGCGATGACATCATG 540
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 569 GGTGCAAAATTTGGCAGCAGCTTCACATCTGCCATACCTTCAGTGGCAGNAGCCATG 628
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 541 TTCCAGCTGGAGTTCTTTATGCCCTCGGCATCATCTTATTTTGTCTCTCAAGATTGTT 600
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 629 TTCTCTCTGGAGTTCTTCTGCTGGCTGGCATCATCTGTCTGCTCAGCCAGAAATTATC 688
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 601 TGGAGCTGAGGGGAGGAGCAGCTGGCCAGACAGGCTCGGATGAAGAAGGGGACCCCG 660
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 689 TGGAGCTGCGGCAGAG---ACAAATGGACCGGCATGCCAAGATCAAGAGAGCCATCAC 745
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 661 TTCCATCATGGTGGTGGCAATTTGTTTCATCACATGCTACTGCCCCAGCGTGTCTGCTAGA 720
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 746 TTCCATCATGGTGGTGGCCATCGTCTTGTTCATCTGCTCTCCAGCGTGGTGTGGCG 805
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 721 CTCATTTCTCTGAGCGGTGCCCTC 746
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 806 ATCCGCATCTTCTGGCTCCTGCACAC 831
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
AL546894 885 bp mRNA linear EST 16-FEB-2001
LOCUS AL546894 LTI_NFL006.PL2 Homo sapiens cdna clone CS0DI026YP01 5
DEFINITION prime, mRNA sequence.
ACCESSION AL546894
VERSION AL546894.1 GI:12880455
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 885)
AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cdna libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
1. .885
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI026YP01"
/tissue_type="LTI_NFL006.PL2"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cdna
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cdna was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : filiang@lifetech.com URL :
http://fulllength.invitrogen.com"
170 a 252 c 229 g 234 t

BASE COUNT
ORIGIN
```

```
Query Match 30.5%; Score 317.2; DB 9; Length 885;
Best Local Similarity 65.0%; Pred. No. 4.8e-76;
Matches 485; Conservative 0; Mismatches 258; Indels 3; Gaps 1;

QY 1 ATGTACAACGGGTCTGTCTGGCGCATCGAGGGGAGCACCATCTCCAGGTGATGCCGCCG 60
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 77 ATAGACAAGAAGAACTGCTGTGTTCGAGATGACTTCATTGTCAAGTGTGGCGCGC 136
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 CTGCTCATTTGGCCCTTTTGTGCTGGCGCAGCTAGGCAATGGGGTCCCGCTGTGTGTTTC 120
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 137 GTGTGGGGCTGGAGTTTATCTTCGGGCTTCTGGGCAATGGCCTTGCCTGTGGATTTC 196
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 TGCCTCCACATGAAGACCTGGAGCCAGCACCTGTTTACCTTTTCAATTTGGCCGTGGCT 180
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 197 TGTTCACACTCAAGTCCCTGGAAATCCAGCCGGATTTTCTGTTCAACCTGGCAGTGGCT 256
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 GATTTCTCTTATGATCTGCCTTTTTCGACAGACTATTACCTCAGACGTAGACAC 240
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 257 GACTTTCTACTGATCATCTGCCTGCCCTTCTTGATGACAACATATGTGAGGCGTTGGGAC 316
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 TGGGCTTTTGGGACATCTCCCTGCCAGTGGGGCTTTCACGTTGGCCATGAACAGGCC 300
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 317 TGGAGTTTGGGACATCCCTTTCGGGCTGATGCTCTTCATGTTGGCTATGAACGCCAG 376
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 GGGAGCATCGTGTCTTACGGTGGTGGCGACAGGTATTTCAAAGTGGTCCACCCC 360
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 377 GGCAGCATCATCTCTCAGGTGGTGGCGTAGACAGGTATTTCGGGGTGGTCCATCCC 436
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 CACCACGGGTGAACACTATCTCCACCGGGTGGCGCTGGCATCTCTGCACCCCTGTGG 420
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 437 CACCACGGCTGAACAAGATCTCCAATCGACAGCAGCCATCATCTCTTGGCTTCTGTGG 496
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 GCCCTGGTCATCTCTGGGAACAGTGTATCTTTTCTGGAGAACCATCTCTGCGCTGCAAGAG 480
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 497 GGCATCACTATTGGCTGACAGTCCACCTCCTGAAGAAGAAGATGCCGATCCAGATGGC 556
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 481 ACGCGCTCTCTCTGTCAGAGCTTCATCATGGAGTGGCGCAATGGCTGGCATGACATCATG 540
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 557 GGTGCAAAATTTGGCAGCAGCTTCAGCATCTGCCATACCTCCAGTGGCAGGAGCCATG 616
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 541 TTCCAGCTGGAGTTCTTTATGCCCTCGGCATCATCTTATTTGCTCTTCAAGATTGTT 600
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 617 TTCTCTCTGGAGTTCTTCTGCTGGCCCTGGGCATCATCTGTTCTGCTCAGCCAGAAATTATC 676
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 601 TGGAGCTGAGGGGAGGAGCAGCTGGCCAGACAGGCTCGGATGAAGAAGGGGACCCCG 660
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 677 TGGAGCTGCGGCAGAG---ACAAATGGACCGGCATGCCAAGATCAAGAGAGCCATCAC 733
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 661 TTCCATCATGGTGGCAATTTGTTTCATCACATGCTACTGCCAGCGTGTCTGCTAGA 720
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 734 TTCCATCATGGTGGCGCATCGTCTTGTTCATCTGCTTCCAGCGGTGTGTGGCG 793
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 721 CTCATTTCTCTGAGCGGTGCCCTC 746
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 794 ATCCGCATCTTCTGGCTCCTGCACAC 819
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
BI837965 801 bp mRNA linear EST 04-OCT-2001
LOCUS 603083801F1 NTH_MGC_120 Homo sapiens cdna clone IMAGE:522688 5',
DEFINITION mRNA sequence.
ACCESSION BI837965
VERSION BI837965.1 GI:15949515
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 801)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
```


Db 331 TGGAGTTGGGGACATCCCTTCCCGCTGATGCTCTTCATGTTGGCTATGAACCCCAA 390
QY 301 GGAGCATCGTGTCTTACGGTGGTGGCTGGACAGTATTTCAAAGTGGTCCACCC 360
Db 391 GGACATCAWCTTCTCAGGTGGTGGCGGTARACAAGTATTTCCGRTGGTCCATCCC 450
QY 361 CACCACGGGTGAACACTATCTCCACCCGGTGGCGCTGGCATCTCTGCAACCCGTGG 420
Db 451 CACCACGGCTGAACAAGAWCTCCAATCGACAGCAGCCATCATCTCTTGGCTTCTGTGG 510
QY 421 GCCTGGTATCTCTGGACAGTATCTTTTCTGGAGAACCATCTCTCGCTGCAAGAG 480
Db 511 GGATCACTATTGGCTGACAGTCCACCTCTGAAGAAGAACGCGATCCAGAATGCG 570
QY 481 ACGGCGCTCTCTGTGAGAGCTTCATCATGGAGTGGCCAAATGGCTGGCATGACATCATG 540
Db 571 GGTGCAATTTGTGACAGCGTTCAGCATCTGCCAWACCTTCCAGTGGCAGAGCCATG 630
QY 541 TTCCAGCTGAGTTCCTTTATGCCCCCTGGGCATCATCTTATTTTGTCTCTTCAAGATTGT 600
Db 631 TTCTCTCTGAGTTCCTCTGCCCCCTGGGCATCAWCTGTCTGCTCAGCCAGAATTATC 690
QY 601 TGAGAGCTGAGGCGGAGGAGCAGCTGGCCAGACAGCTCGGATGAAGAGGCGACCCGG 660
Db 691 TGAGGCTGGGCGAGAG--ACAAATGGAGCGGCGCATGCAAGATCAAGAGAGCCACACC 747
QY 661 TTCATCATGCTGCTGGCAATTTGTTCATCACATGCTACTGCCACGCTGTCTGCTAGA 720
Db 748 TTCATCATGCTGCTGGCAWCTTTGTCAWCTGCTCTCTCCAGCGTGGTGTGGGAT 807
QY 721 CTCTATTTC 729
Db 808 CGCATCTTC 816

RESULT 6
BI820995
LOCUS 603035649F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5176597 5',
DEFINITION mRNA sequence.
ACCESSION BI820995
VERSION BI820995.1 GI:15932545
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 876)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11440 row: a column: 14
High quality sequence start: 6
High quality sequence stop: 847.
Location/Qualifiers
1. .876
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5176597"
/lab_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1

FEATURES
source

male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

BASE COUNT 168 a 247 c 230 g 231 t
ORIGIN

Query Match 26.8%; Score 278.6; DB 10; Length 876;
Best Local Similarity 64.6%; Pred. No. 1.9e-65;
Matches 463; Conservative 0; Mismatches 249; Indels 5; Gaps 3;

QY 1 ATGTACAACGGGTGCTGCGCATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 60
Db 91 ATAGACAAGAAGACTGCTGTGTTCCGAGATGACTTCATTTCAAGGTGTTGCCGCCG 150
QY 61 CTGCTCATTTTGGCCCTTTGTGCTGGCGGCACATAGGCAATGGGGTCCGCCCTGTGTGTTTC 120
Db 151 GTGTTGGGGCTGGAGTTTATCTTCGGGCTTCTGGGCAATGGCCCTTGCCTGTGGATTTTC 210
QY 121 TGCCTCCACATGAGACCTGGAGCCCGCAGCTGTTTACCTTTTCAATTTGGCCGTGGCT 180
Db 211 TGTTCACCTCAAGTCTTGGAAATCCAGCCGGATTTCTCTGTTCACCTGGCAGTGGCT 270
QY 181 GATTTCTCTTATGATCTGCCCTGCTTTTCGGACAGACTATTACCTCAGACCTAGACAC 240
Db 271 GACTTCTACTGATCATCTGCCCTTCCCTGATGACACACTATGTGAGGCTTGGGAC 330
QY 241 TGGGCTTTTGGGACATTCCTTCCGAGTGGGCTCTTACGTTGGCCATGAACAGGGCC 300
Db 331 TGGAACTTTGGGACATCCCTTCCGCGCTGATGCTCTTCATGTTGGCTATGAAC-CGCCA 389
QY 301 GGGAGCATCGTGTTCCTTAGCGTGGTGGCTGGCGACAGGATTTTCAAAGTGGTCCACCCC 360
Db 390 GCGAGCATCATCTCTCCACGCTGCTGAGAGAGTATTTTCCGGGTGGTCCATCCC 449
QY 361 CACACGCGGTGAACACTATCTCCACCGGTGGCGGCTGGCATCTCTGACACCCCTGTGG 420
Db 450 CACCACGCCCTGAACAAGATCTCCAATCGGACAGAGCCATCATCTCTTGGCTTCTGTGG 509
QY 421 GCCTGTGTCATCTCTGGGAACAGTGTATCTTTTCTGGAGAACCATCTCTCGCTGCAAGAG 480
Db 510 GGCATCACTATTGGCCCTGACAGTCCACCTCTCGAAGAAGAAGATGCCGATCCAGAATGCG 569
QY 481 ACGGCGCTCTCTGTGAGAGCTTCATCATGGAGTGGCCCAATGGCTGGCATGACATCATG 540
Db 570 GGTGCAAAATTTGTGACAGCAGCTTCAGCATCTGCCATACCTTCCAGTGGCAGAGCCATG 629
QY 541 TTCCAGCTGGAGTCTTTATGCCCCCTCGGCATCATCTTATTTTGTCTCTTCAAGATTGTT 600
Db 630 TTCTCTCTGGAGTCTTCTTCCGCCCTTGGGCATCATCTCTGTTCTGCTCAGCCAGAAATATC 689
QY 601 TGGAGCTGAGGCGGAGCAGCAGCTGGCCAGACAGCTCGGATGAAGAGGCGACCCCG 660
Db 690 TGGAGCTCGGCGCAGAGACA--AATGGACGGCATTTGCCAAGATCAAGAGAGCCATCACC 747
QY 661 TTCATCATGCT--GGTGGCAATTTGTTCATCATGCTTACCTGCCAGCGGTGTCTG 715
Db 748 TTCATCATGCTTGGTGGCCATCGTCTTTTGTCTATCTTCTTCCCTTCCAGCGTGGTTG 804

RESULT 7
BI908162
LOCUS 603067472F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5216512 5',
DEFINITION mRNA sequence.
ACCESSION BI908162
VERSION BI908162.1 GI:16171076
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 734)

NH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L14M11543 row: p column: 17

High quality sequence stop: 731.

FEATURES

Location/Qualifiers

1..734

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5216512"

/clone_lib="NIH_MGC_118"

/tissue_type="leukocyte"

/lab_host="DH10B"

/note="vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV

(destroyed); RNA source leukocytes from anonymous pool of

non-activated adult donors. Library is oligo-dT primed

and directionally cloned (EcoRV site is destroyed upon

cloning). Average insert size 1.7 kb, insert size range

1.2-3.3 kb. Library is normalized and enriched for

full-length clones and was constructed by C. Gruber

(Invitrogen). Research Genetics tracking code 027. Note:

this is a NIH_MGC Library."

149 a 203 c 190 g 192 t

BASE COUNT

Query Match 26.0%; Score 270.4; DB 10; Length 734;

Best Local Similarity 65.8%; Pred. No. 3.1e-63;

Matches 408; Conservative 0; Mismatches 211; Indels 1; Gaps 1;

Qy 1 ATGTACACGGGTGCTGCTGCGCATCGAGGGGACACCATCTCCAGGTGATCGCGCG 60

Db 85 ATAGACAAGAAGAACTGCTGTGTCGAGATGACATTCATGTCGAAGGTGTTCGCCCG 144

Qy 61 CTGCTCATTTGGGCTTGTGCTGGCGCACTAGGCAATGGGTGCGCTGTGTGGTTTC 120

Db 145 GTGTTGGGCTGGAGTTATCTTTGGGCTTCTGGCAATGGCCTTGGCCTGTGATTTTC 204

Qy 121 TGCTTCCACATGAAGACCTGGAAGCCAGCACACTGTTTACCTTTTCAATTTGGCGGTGGCT 180

Db 205 TGTGTTCCACCTCAAGTCTCTGGAATCCAGCGGATTTTCTGTTTCAACCTGGCAGTAGCT 264

Qy 181 GATTTCTCTTATGATCTGCTGCTGCTTTCCGACACACTATTACCTCAGACGTAGACAC 240

Db 265 GACTTTCTACTGATCATCTGCTGCTGCTGCTGATGACTACTATGTGGCGCTTCAGAC 324

Qy 241 TGGGCTTTGGGACATTCCTCGGAGTGGGCTTTCACGTTGGCATGAACAGGCGC 300

Db 325 TGGAGTTTGGGACATTCCTTGGCGCTGTGCTCTTCATGTTGCAATGAAC-CGCCA 383

Qy 301 GGGAGCATCGGTGTTCTTACGGTGGCTGCGGACAGGATTTTCAAAGTGGTCCACCCC 360

Db 384 GGCAGCATCATATTCCTCACGGTGTGGCGGTAGACAGGATTTTCCGGGTGGTCCATCCC 443

Qy 361 CACCAGGGGTGAACACTATCTCCACCGGGGTGGCGGTGGCATCGTGTGACCCCTGTGG 420

Db 444 CACCAGCCCTGAACAAGATTCCTCAATTTGGACAGCAGCCATCTCTTGCTTCTGTGG 503

Qy 421 GCCCTGCTCATCTGGGAACAGTCTATCTTTTGTGCTGGAGACCATCTCTCGTGCAGAG 480

Db 504 GGCATCACTGTTGGCCTAAACAGTCCACCTCTCTGAAGAAGAAGTTGCTGATCCAGAATGGC 563

Qy 481 ACGCCCTCTCTCTGTGAGAGCTTCATCATGAGTGGCCCAATGGCTGGCATGACATCATG 540

Db 564 ACTCAAATGTGTGCATCAGCTTCAGCATCTGCCATACCTTCGGTGGCAGCAAGCTATG 623

Qy 541 TTCAGCTGAGGTCTTTATGCCCCCTGGCATCATCTTATTTGCTCTCTTCAAGATGTTT 600

Db 624 TTCTCTCTGGAGTCTTCTCTGCCCTGGCATCATCTGTTCTGCTCAGCAGAAATATC 683

Qy 601 TGGAGCTGAGGCGGAGGACA 620

Db 684 TGGAGCTGCGGCAGACACA 703

RESULT 8

BM008116

LOCUS 603618002F1 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:5450832 5',

DEFINITION mRNA sequence.

ACCESSION BM008116

VERSION BM008116.1 GI:16522470

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 752)

NH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Dr. Mark Watson

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1CM1940 row: d column: 01

High quality sequence stop: 752.

FEATURES

Location/Qualifiers

1..752

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5450832"

/clone_lib="NIH_MGC_113"

/lab_host="DH10B (phage-resistant)"

/note="Organ: spleen; Vector: pOTB7; Site_1: XhoI; Site_2:

EcoRI; cDNA made by oligo-dT priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:

GGCAGCAG(G). Library constructed by Ling Hong in the

Laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a

NIH_MGC Library."

153 a 216 c 193 g 190 t

BASE COUNT

Query Match 25.7%; Score 267.2; DB 10; Length 752;

Best Local Similarity 65.5%; Pred. No. 2.4e-62;

Matches 406; Conservative 0; Mismatches 213; Indels 1; Gaps 1;

Qy 1 ATGTACACGGGTGCTGCTGCGCATCGAGGGGACACCATCTCCAGGTGATCGCGCG 60

Db 130 ATACACAAGAAGAACTGCTGTGTTCCGAGATGACATTCATGTCGAAGGTGTGCGCGCG 189

Qy 61 CTGCTCATTTGGCCTTTGTGCTGGCGCACTAGGCAATGGGTGCGCCTGTGTGGTTTC 120

Db 190 GTGTTGGGCTGGAGTTTATCTTCGGGCTTCTGCGCAATGGCCTTGGCCTGTGGATTTTC 249

Qy 121 TGCTTCCACATGAAGACCTGGAAGCCAGCACACTGTTTACCTTTTCAATTTGGCGGTGGCT 180

Db 250 TGTGTTCCACCTCAAGTCTCTGGAATCCAGCGGATTTTCTGTTTCAACCTGGCAGTGGCT 309

D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
e mouse tissues.

FEATURES
source

1. 660
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="483341H22"
/clone_lib="RIKEN full-length enriched, 0 day neonate head"
/sex="mixed"
/tissue_type="head"
/dev_stage="0 day neonate"
/lab_host="DH10B"
/note="Site.1: Sali; Site.2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'
GAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCGATTAAATTAATCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 131 a 162 g 184 t 1 others
ORIGINQuery Match 24.5%; Score 254.6; DB 9; Length 660;
Best Local Similarity 63.7%; Pred. No. 6.5e-59;

Matches 386; Conservative 0; Mismatches 220; Indels 0; Gaps 0;
QY 1 ATGTACAAACGGTCTGCTGCGCATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 60
DB 54 ATAAACGGCAAGAACTCTGTGTGTTCCGAGATGAACAACATGCCAAGGTCTTGCCACCG 113
QY 61 CTGCTCATTTGGCCCTTTGTGTGGGCGCACTAGGCAATGGGTCGCCCTGTGTGGTTTC 120
DB 114 GTGTTGGGGCTGGAATTTGTTCGGACTCTCTGGCAATGGCCCTTGCTTGTGATTTTC 173
QY 121 TGCTTCCACATGAAGACTGGAAGCCAGCACTGTTTACCTTTTCAATTTGGCGTGGCT 180
DB 174 TGTTCACCTCAAGTCTCGAAATCCAGCGGATTTTCTGTTCACATTTGGCGTGGCT 233
QY 181 GATTTCCTCTTATGATCTGCTGCTCTTCGGACAGACTATTACCTCAGACGTAGACAC 240
DB 234 GACTTCTCTGATCATCTGCTGCTGCTCTGAGCGACAACATATGTCCTAACTGGGAC 293
QY 241 TGGGCTTTTGGGACATTCCTTCCCGAGTGGGCTCTTACGTTGGCCATGAACAGGCC 300
DB 294 TGGAGTTTCGGAGCATCCCTGCGGTGTGATGCTCTTATGTTGGCTATGAACCGACAG 353
QY 301 GGGAGCATGCTGTTCCTTACGGTGGTGGCTGCGGACAGTATTTCAAAGTGGTCCACCC 360
DB 354 GCGAGCATCATCTTCTCCACCGTGGTGGCTGAGCGCTACTTCCGGGTGGTCCATCCA 413
QY 361 CACCACCGGTGAACACTATCTCCACCGGTGGCTGGCTGGCATCGTCTGCACCTCTGG 420
DB 414 CACCCTTTGAACAGATCTCCACCGGACGGCGCCCATCATTTCTTCTTCTTGG 473
QY 421 GCGCTGTGTCATCTGGGAACAGTATCTTTTGTGTGAGAACCATCTCTCGTGTGCAAGAG 480
DB 474 GGTCTACCATCGCTGCTGCTGCTCTCTATACAAACATGATGACCAAAATGGC 533
QY 481 ACGGCGCTCTCTGTGAGAGCTTCATCATGAGTCGCGCATGCTGGCATGATCATG 540
DB 534 GAGGCATATCTGTGTAGAGCTTCAGAGCTTCAGCACTGTTACAACTTCAGGTGCGCATG 593
QY 541 TTCAGCTGGAGTCTTTTATGCCCTCGGCATCATCTTATTTTCTCTTCAAGATGTT 600
DB 594 TTCTCTTTGGAAATCTTCTTGGCCCTGGCCATCATCTTGTCTGCTCANGCAGGATCATC 653
QY 601 TGGAGC 606
DB 654 TGGAGC 659

RESULT 11
BB873597
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BB873597 367 bp mRNA linear EST 27-NOV-2001
BB873597 RIKEN full-length enriched, 15 days embryo brain Mus
musculus cDNA clone G630048L01 5', mRNA sequence.

BB873597
EST.
BB873597.1 GI:17119807

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 367)

Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K., Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akashira, S., Tanaka, T., Tomaru, A., Toya, T., Watahiki, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)

Unpublished (2001)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.

FEATURES
source
1. 367
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G630048L01"
/clone_lib="RIKEN full-length enriched, 15 days embryo
brain"
/tissue_type="brain"
/dev_stage="15 days embryo"
BASE COUNT 65 a 106 c 92 g 104 t
ORIGIN

Query Match 24.3%; Score 253.2; DB 9; Length 367;
Best Local Similarity 83.2%; Pred. No. 1.2e-58;
Matches 288; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 ATGTACAAACGGGTGCTGCTCGGCATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 60
Db 22 ATGACAAACGGGTGCTGCTCATCGAGGGGAGGCCATCTCCCAAGTGTATGCTCCT 81

QY 61 CTGCTCATTTGTCGCTTTGTCGGGCGCACTAGGCAATGGGTGCGCCCTGTGTGTTTC 120
Db 82 CTACTCATCTGCTGCTGCTGCTGGCGCCCTGGGCAACGGCATAGCCCTGTGCGGCTC 141

QY 121 TGCTTCACATGAAGACCTGGAAGCCAGACACTGTATACCTTTCAATTTGGCGGTGCT 180
Db 142 TGCTTCACATGAAGACCTGGAAGTCAAGCACTATTTACCTTTTCACTTGGCTGTGCC 201

QY 181 GATTTCTCTTATGATCTGCTGCTGCTTTTCGGACAGACATTAACCTCAGACAGTACAC 240
Db 202 GATTTCTCTCATGATCTGCTTACCCCTTCGGACAGACACTACTACCTCAGACAGACAC 261

QY 241 TGGCTTTTGGGACATCCCTGCGGAGTGGGCTCTTACGTTGGCCATGAACAGGCC 300
Db 262 TGGATTTTGGAGATACGCTGCTGCTGCTGCTCTCTCTCAAGCTGGACATGAATAGGCC 321

QY 301 GGGAGCATCTGTTCCTTACGTTGGTGGTTCGGACAGAGTATTTCA 346
Db 322 GGGAGCATCTGTTCCTTACGTTGGTGGTGTGGATAGGATTTTCA 367

RESULT 12
BB849390 422 bp mRNA linear EST 26-NOV-2001
LOCUS BB849390
DEFINITION BB849390 RIKEN full-length enriched, adult inner ear Mus musculus

cDNA clone F930010D18 5', mRNA sequence.
BB849390
VERSION BB849390.1 GI:17090844
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
1 (bases 1 to 422)
Akimura,T., Atakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Ishii
Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuuyama,T.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,
Saito,K., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.,
Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
2001)

TITLE
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.

FEATURES
source
1. 422
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="F930010D18"
/clone_lib="RIKEN full-length enriched, adult inner ear"
/tissue_type="inner ear"
/dev_stage="adult"
BASE COUNT 80 a 109 c 119 g 114 t
ORIGIN

Query Match 22.1%; Score 230; DB 9; Length 422;
Best Local Similarity 77.5%; Pred. No. 3e-52;
Matches 330; Conservative 0; Mismatches 90; Indels 6; Gaps 4;

QY 193 ATGATCTGCTGCTGCTTTTCGGACAGACTA-TTACCTCAGACGTAGACACTGGCTTTGG 251
Db 2 ATGATCTGCTTACCCCTTCGGACAGACTATACCTCAGACGACACTGATTTTGG 61

QY 252 GGACATTCCTCGGAGTGGGCTCTTCACGTTGCCATGAACAGGCCGGGAGCATCGT 311
Db 62 AGATATCGCTGTGCTGCTGCTCTTCAAGCTGCCATGAATAGGG-CGGAGCATGTG 120

```
Qy 312 GTTCTTACGGTGGTGGCGGACAGTATTTCAAAGTGTCCACCCACCGCGGT 371
Db 121 TTCTCTACTGTGGTGGTGTAGTAGTATTTCAAAGTGTCCACCCACCGTGT 180
Qy 372 GAACACTATCTCCACCGGCTGGCGGTGGCATCGTCTGCACCCCTGTGGCCCTGGTCA 431
Db 181 GAATGCCATCTCCACCGGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
Qy 432 CTTGGAAACAGTATCTTTTCTGGGAGAACCATCTCTGCGTGGCAAGAGCGGCGTCTC 491
Db 241 CTTGGGAGTGTATCTCTGTATGGAGAGTACCTGTGTGGAGGAGACATGTCTCTC 300
Qy 492 CTGTGAGAGCTTCATCATGAGTGGCGCAATGGCTGGCATGACATCATGTTCCAGCTGGA 551
Db 301 CTGTGAGAGCTTAATAATGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
Qy 552 GTTCTTTATGCGGCTGGCGCATCTTATTTGTTGCTCTTCAAGATGTTTGGAGCCTGAG 611
Db 361 GTCTTCT---GCCCTGACAATCATCTTGTCTGCT-GGTCAAGGTCTGTTGGAGCCTGAG 416
Qy 612 GCGGAG 617
Db 417 ACGGAG 422

RESULT 13
BF074336 550 bp mRNA linear EST 25-APR-2001
LOCUS 221723 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION BF074336
VERSION BF074336.1 GI:10867847
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 550)
Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,
Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett
, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-Mckown, C.G.,
Pertes, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and
Keele, J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smithemail.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCAGTCACGACG
Plate: 80 row: K column: 23
Seq primer: ATTAGTGACACTATAG.
Location/Qualifiers
1..550
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 2BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from testis, thymus,
semitendinosus muscle, longissimus muscle, pancreas,
adrenal, and endometrium."
97 a 164 c 153 g 136 t
"

FEATURES
source
1..550
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 2BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from testis, thymus,
semitendinosus muscle, longissimus muscle, pancreas,
adrenal, and endometrium."
BASE COUNT
97 a 164 c 153 g 136 t
"
```

```
ORIGIN
Query Match 21.0%; Score 218.2; DB 10; Length 550;
Best Local Similarity 68.6%; Pred. No. 5.9e-49;
Matches 301; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

Qy 16 TGTCTGCCCATCGAGGGGACACCATCTCCAGAGTGATGCCCGCTGCTCATTTGTGGCC 75
Db 92 TGTCTGTGTTCGGGAGTACTTATTCGCAAGTGTCTGCCCGGCTGCTGGGCTGGAG 151
Qy 76 TTTGTCTGGGCGGACTAGGCAATGGGTGCGCTGTGTGGTTTCTGCTCCACATGAAG 135
Db 152 TTTGTCTGGGCGGCTGCGGCAATGGGTGCGCTGTGTGGTTTCTGCTCCACATGAAG 211
Qy 136 ACCTGGAGCCGACGACTGTTTACCTTTTCAATTGGCGGTGGCTGATTTCTCTCTATG 195
Db 212 TCTGGGAAGCGCGGGTTCCTGTTCACTTTGGCGGTGGCTGACTTCTCTCTGATC 271
Qy 196 ATCTGCTGCTGCTTTTCGGACAGACTATTACCTCAGACCTAGACACTGGGCTTTTGGGAC 255
Db 272 ATCTGCTGCTGCTTTTCGGCGGACACTACCTGCGGAGTGGGACTGGAAGTTTGGGAG 331
Qy 256 ATTCCTGCGGAGTGGGCTCTTACGTTGGCCATGAACAGGCGGAGAGCATCGTGTTC 315
Db 332 ATCCCTGCGGCTCATGCTCTTCAATGTTGGCCATGAACCGCCAGGCGAGCATCATTTTC 391
Qy 316 CTTACGGTGGTGGGCGGACAGGTATTTCAAAGTGTGCCACCCACCGCGGTGAAC 375
Db 392 CTCACCGTGGTGGGCGGTGATAGGTACTTCGGGTGGTTCACCCACCATGCGCTGAAC 451
Qy 376 ACTATCTCCACCGGCTGGCGGCTGGCATCTCTGCACCTGTGGCGCTGCTGTCATCCTG 435
Db 452 AAGATCTCCATCGGACAGCGGCGCATCATCTCTGCTCTTGTGGGCGATCACCATCGCG 511
Qy 436 GGAACAGTGTATCTTTTGC 454
Db 512 CTGACGGTCCATCTCCCTAC 530

RESULT 14
BF489649 636 bp mRNA linear EST 28-AUG-2001
LOCUS 603032118F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5173205 5',
DEFINITION BF489649
VERSION BF489649.1 GI:15328877
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 636)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaps@email.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LIA01431 row: d column: 06
High quality sequence stop: 634.
Location/Qualifiers
1..636
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:5173205"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
"
```


/note="Organ: pooled brain, lung, testis; Vector: pcMV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dr primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH/MGC Library."

BASE COUNT 124 a 177 c 165 g 170 t

Query Match 20.7%; Score 215.8; DB 10; Length 636;
Best Local Similarity 64.5%; Pred. No. 2.9e-48;
Matches 322; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

QY 43 TCCAGGTGATGCGCGCTGCTCATTTGGCTTTGGTGTGGGCGACTAGGCAATGGG 102
DB 137 TGGCAAGGTGCGCGCGGTGTTGGGCTGGAGTTTATCTTTGGGCTTCTGGGCAATGGC 196
QY 103 GTCCCTGTGTGTTCTCTCTCCACATGAAGACCTGGAAGCCGACGACTGTTTACCTT 162
DB 197 CTGGCTGTGGATTTCTGTTTCCACCTCAAGTCTGGAATCCAGCCGGATTTTCTCTG 256
QY 163 TTCAATTTGGCGCTGATTTCTCTCTATGATCTGCTGCTTTTTCGGACAGACTAT 222
DB 257 TTCAACTGGCAGTAGCTGACTTTCTACTGATCATCTGCTGCGTTCGTGATGGACTAC 316
QY 223 TACCTCAGAGTACAGCTGGGCTTTTGGGCGATTCCTTCCGAGTGGGGCTCTTCAAG 282
DB 317 TATGTGGGGCTTACAGCTGGAATTTTGGGAGATCCCTTGGCGGCTGGTGTCTTTCATG 376
QY 283 TTGGCATGAACAGGCGCGGAGCATGCTTCTTACGCTGTGGCTGCGGACAGTAT 342
DB 377 TTGGCATGAACCGCAGGCGAGCATCATCTTCTTACGCTGTGGCTGCGGAGTAGAGTAT 436
QY 343 TTCAAGTGTCTCCACCCCGGCGGTGAACACTATCTCCACCGCGGTGGGGCTGGC 402
DB 437 TTCGGGTGTCTCCACCCCGGCGGTGAACACTATCTCCACCGCGGTGGGGCTGGC 496
QY 403 ATGCTGTGACCTGTGGGCGCTGGTATCTCTGGGAACAGTATCTTTTGTGGAGAAC 462
DB 497 ATCTCTTGGCTTGTGGGCGATCAGTGTGGCTTACAGTCCACCTCTTGAAGAAG 556
QY 463 CATCTCTGCTGCAAGAGCGCGCTCTCTGTGAGAGCTTCATCATGAGTGGGCAAT 522
DB 557 TTGCTGTGAGATGGCCCTGCAATGTGTGATCATGCTTACAGTCTGCCATACCTTC 616
QY 523 GGCTGGCATGACATGATGT 541
DB 617 CGGTGGCAGCAAGTATGT 635

RESULT 15
BF230383 532 bp mRNA linear EST 14-NOV-2000
LOCUS 251977 BARC 5BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION BF230383
ACCESSION BF230383
VERSION BF230383.1 GI:11168302
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 532)
AUTHORS Sonstegard,T.S., Capuco,A.V., Van Tassel,C.P., Ashwell,M.S. and Wells,K.D.
TITLE Mapping of Expressed Sequence Tags from a normalized bovine mammary gland cDNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Sonstegard ts

USDA, ARS, Beltsville Agricultural Research Center
Bldg. 200 Rm 2A, Beltsville, MD 20705, USA
Tel: 301 504 8416
Fax: 301 504 8414
Email: tadselpsi@barc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACACGCTATGACCAT
BACKWARD: GTTTCCCGAGTCACGACG
Plate: 110 row: C column: 22
Seq primer: ATTAGGTGACACTATAG.

FEATURES
Source
1. .532
Location/Qualifiers
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="BARC 5BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/notes="Vector: pcMV SPORT6; Site.1: XbaI; Site.2: XhoI;
Library made from pooled mRNA isolated from mammary
tissues at eight physiological, developmental, and disease
states."

BASE COUNT 90 a 163 c 145 g 134 t

Query Match 20.5%; Score 213.8; DB 10; Length 532;
Best Local Similarity 69.5%; Pred. No. 9.4e-48;
Matches 290; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 16 TGTGTGCGCATCGAGGGGACACATCTCCAGGTGATCGCGCGCTGCTCATTTGTGGCC 75
DB 109 TGTGTGTGTCGCGATGACTTCATTTGCCAAGTGTGCGCGCGCTGCTGGGGCTGGAG 168
QY 76 TTTGTGTGCGGCGACTAGGCAATGGGGTGCCTCTGTGTGTTTCTCTTCCACATGAAG 135
DB 169 TTGCTGTGTCGGGCTCCTGGGCAATGGCCTTGGCCTGTGGATCTTCTGCTTCCACCTCAAG 228
QY 136 ACCTGGAAGCCCGAGCACTGTTTACCTTTTCAATTTGGCCGTGGCTGATTTCCCTCTTATG 195
DB 229 TCCTGGAAAGCCCGCGGTTTTCCTGTTCAACTTGGCCGTGGCTGACTTTCCTCTGATC 288
QY 196 ATCTGCTGCTTTTTCGGACAGACTATTACCTCAGACGTAGACACTGGGCTTTTGGGAC 255
DB 289 ATCTGCTGCGCTTCTCTGCGGACACTAGCTGCGGAGTGGGACTGGAAGTTTGGGAG 348
QY 256 ATTCCCTGCGGAGTGGGCTCTTACCTTGGCCATGAACAGGCGCGGAGCATGCTGTTTC 315
DB 349 ATCCCTTGGCGGCTCATGCTCTTTCATGTTGGCCATGAACGCCAGGCGACGATCATTTTC 408
QY 316 CTTACGGTGTGGTGGGACAGTATTTCAAAGTGTTCACCCCGCCACCGCGGTGAAC 375
DB 409 CTCACCGTGGTGGCGGTGGATAGTACTTCGGGTGCTGCTATCCCCACCATGCCCTGAAC 468
QY 376 ACTATCTCCACCGGCTGGGCTGGCATGCTGTGACCTGTGGGCGCTGGTCAATC 432
DB 469 AAGATCTCCAATCGGACAGCGGGCATCATCTCTGCTGCTTGTGGGCGATCACCATC 525

Search completed: October 29, 2002, 03:44:18
Job time : 1736 secs

THIS PAGE BLANK (US)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2002, 00:15:03 ; Search time 50 Seconds
(without alignments)
5114.090 Million cell updates/sec

Title: US-09-886-041-1
Perfect score: 1041
Sequence: 1 atgtacaacggctgctg.....acattgttgagtgcaactga 1041

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCITUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	229.6	22.1	545	1 US-07-915-966C-1	Sequence 1, Appl
2	229.6	22.1	545	2 US-08-771-182-1	Sequence 1, Appl
3	229.6	22.1	545	3 US-08-853-194-1	Sequence 1, Appl
4	188.8	18.1	1594	2 US-08-955-713-1	Sequence 1, Appl
5	188	18.1	1435	2 US-08-955-713-3	Sequence 3, Appl
6	139.2	13.4	960	3 US-09-130-749-1	Sequence 1, Appl
7	139.2	13.4	960	3 US-09-130-749-1	Sequence 1, Appl
8	109	10.5	1901	1 US-08-153-848-43	Sequence 43, Appl
9	109	10.5	1901	3 US-09-299-843A-43	Sequence 43, Appl
10	109	10.5	1901	4 US-09-088-337B-43	Sequence 43, Appl
11	109	10.5	1901	5 PCT-US93-11153-43	Sequence 43, Appl
12	109	10.5	2453	5 PCT-US95-07180-1	Sequence 1, Appl
13	104.2	10.0	1597	2 US-08-724-974A-1	Sequence 1, Appl
14	92.6	8.9	1842	1 US-08-442-134A-1	Sequence 1, Appl
15	92.6	8.9	1842	1 US-08-444-581B-1	Sequence 1, Appl
16	92.6	8.9	1842	1 US-08-446-088A-1	Sequence 1, Appl
17	81.4	7.8	1378	1 US-08-759-848-2	Sequence 2, Appl
18	81.4	7.8	1378	5 PCT-US95-09383-2	Sequence 2, Appl
19	78.4	7.5	1155	3 US-09-053-866-3	Sequence 3, Appl
20	78.2	7.5	1101	1 US-08-759-848-7	Sequence 7, Appl
21	78.2	7.5	1101	5 PCT-US95-09383-7	Sequence 7, Appl
22	78.2	7.5	3680	1 US-08-759-848-1	Sequence 1, Appl
23	78.2	7.5	3680	5 PCT-US95-09383-1	Sequence 1, Appl
24	75.8	7.3	1255	1 US-08-097-938-3	Sequence 3, Appl
25	75.8	7.3	1255	1 US-08-476-000-3	Sequence 3, Appl
26	75.8	7.3	1255	1 US-08-472-840-3	Sequence 3, Appl
27	75.8	7.3	1255	2 US-08-476-976-3	Sequence 3, Appl

28	75.8	7.3	1255	3 US-08-474-410-3	Sequence 3, Appl
29	75.8	7.3	1255	4 US-08-486-673B-3	Sequence 3, Appl
30	72.8	7.0	1265	1 US-07-816-283-3	Sequence 3, Appl
31	72.8	7.0	1265	1 US-08-417-103-3	Sequence 3, Appl
32	72.6	7.0	1414	1 US-08-476-000-62	Sequence 62, Appl
33	72.6	7.0	1414	1 US-08-472-840-62	Sequence 62, Appl
34	72.6	7.0	1414	2 US-08-476-976-62	Sequence 62, Appl
35	72.6	7.0	1414	3 US-08-474-410-62	Sequence 62, Appl
36	72.6	7.0	1414	4 US-08-486-673B-62	Sequence 62, Appl
37	69.6	6.7	1296	1 US-07-816-283-9	Sequence 9, Appl
38	69.6	6.7	1296	1 US-08-417-103-9	Sequence 9, Appl
39	68.8	6.6	1475	1 US-08-097-938-1	Sequence 1, Appl
40	68.8	6.6	1475	1 US-08-476-000-1	Sequence 1, Appl
41	68.8	6.6	1475	1 US-08-472-840-1	Sequence 1, Appl
42	68.8	6.6	1475	2 US-08-476-976-1	Sequence 1, Appl
43	68.8	6.6	1475	3 US-08-474-410-1	Sequence 1, Appl
44	68.8	6.6	1475	4 US-08-486-673B-1	Sequence 1, Appl
45	68.8	6.6	2884	1 US-08-148-209A-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-07-915-966C-1
; Sequence 1, Application US/07915966C
; Patent No. 5668006
; GENERAL INFORMATION:
; APPLICANT: Hadcock Dr., John R.
; APPLICANT: Ozenberger Dr., Bradley A.
; APPLICANT: Pausch Dr., Mark H.
; TITLE OF INVENTION: Receptor Identification Method
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Home Products Corporation
; STREET: One Campus Drive
; CITY: Parsippany
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/915,966C
; FILING DATE: 17-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthews, Gale M.
; REGISTRATION NUMBER: 32,269
; REFERENCE/DOCKET NUMBER: 31,829-00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-683-2134
; TELEFAX: 201-683-4117
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 545 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Rat
; PUBLICATION INFORMATION:
; AUTHORS: Hadcock Dr., John R.
; AUTHORS: Dr. Ozenberger, Bradley A.
; AUTHORS: Dr. Pausch, Mark H.
; TITLE: Receptor Identification Method
; DATE: 17-JUL-1992

Sequence 1, Application US/08853194
Patent No. 6077666
GENERAL INFORMATION:
APPLICANT: Hadcock Dr., John R.
APPLICANT: Ozenberger Dr., Bradley A.
APPLICANT: Pausch Dr., Mark H.
TITLE OF INVENTION: Receptor Identification Method
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: NJ
COUNTRY: United States of America
ZIP: 06904-0060
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/853,194
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/915,966
FILING DATE: 17-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Tsevdos Dr., Estelle J.
REGISTRATION NUMBER: 31,145
REFERENCE/DOCKET NUMBER: 31829-00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 203-321-2361
TELEFAX: 203-321-2971
TELEX: 710-474-4059
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 545 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Rat
PUBLICATION INFORMATION:
AUTHORS: Hadcock Dr., John R.
AUTHORS: Dr. Ozenberger, Bradley A.
AUTHORS: Dr. Pausch, Mark
TITLE: Receptor Identification Method
DATE: 17-JUL-1992
US-08-853-194-1

Query Match 22.1%; Score 229.6; DB 3; Length 545;
Best Local Similarity 65.0%; Pred. No. 4.5e-57;
Matches 356; Conservative 0; Mismatches 189; Indels 3; Gaps 1;
QY 157 TACCTTTTCAATTTGGCCGCTGATTTCCCTTATGATCTGCTTTCGGACA 216
DB 1 TTCTGGTGAACCTGGTGGGCTGACTTTCTCTGATCATTTGCTGCTTTGAGC 60
QY 217 GACTATTACCTAGACGTAGACCTGGCTTTTGGGACATCCCTGCGGAGTGGGGCTC 276
DB 61 GACAACATGTTCAGAACCTGGGACTGGAGGTTCGGGAGCATCCCTGCGGCTGTGCTC 120
QY 277 TTCACGTGGCCATGAACAGGGCCGAGCATGCTTCTTACGGTGGTGGCTGGGAC 336
DB 121 TTCATGTTGGCCATGAACAGGACAGGACGATCATCTTCTACGGTGGTGGCTGGAC 180
QY 337 AGGTATTTCAAGTGTCCACCCACCACCGGGTGAACACTATCTCCACCCCGGGTGGC 396
DB 181 AGGTACTTCAGGTGTCTCCACCCGACCACTTCTCTGAACAAGATCTCCAACCGGACGGC 240

QY 397 GCTGGCATCGTCTGCACCCCTGTGGCCCTGGTGCATCCTCGGAACAGTGTATCTTTTGGTGC 456
DB 241 GCATCATCTCTTCTTCTTCTGGGCGCATCACCATCGGCTGACAGTCCACCTCCTCTAC 300
QY 457 GAGAACCATCTCTGCGTGCAAGAGAGCGGCGTCTCCTGTGAGAGCTTCATCATGAGTGC 516
DB 301 ACGGACATGATACCCGAAACGCGGATGCAAAACCTGTGCAGCAGTTTGTAGCATCTGCTAC 360
QY 517 GCCAATGGCTGCATGACATCATGTTCCAGCTGGAGTTCTTTATGCCCTCGGCATCATC 576
DB 361 ACTTTCAGGTGGCAGCATGCAATGTTCCCTCTTGGAAATCTCTCTGCCCTGGGCATCATC 420
QY 577 TTATTTTGTCTCTCAAGATTGTTTGGAGCTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 636
DB 421 CTGTTCTGCTCTGGCAGGATCATTTGGAGGCTTAAGGACAGAG---ACAGATGGACAGGCAC 477
QY 637 GCTCGGATGAAGAAGCGGACCGGTTTCATCATGTTGGTGGTGGCAATTTGTTCATCATATGC 696
DB 478 GTCAAGATCAAGAGGCGCATCAACTTCATCATGTTGGTGGTGGCAATTTGTTCATCATATGC 537
QY 697 TACCTGCC 704
DB 538 TGGCTGCC 545
RESULT 4
US-08-955-713-1
Sequence 1, Application US/08955713
Patent No. 5955308
GENERAL INFORMATION:
APPLICANT: SATHE, GANESH
APPLICANT: MOONEY, JEFFREY
APPLICANT: BERGSMAN, DEREK
APPLICANT: HALSEY, WENDY
TITLE OF INVENTION: CDNA CLONE HEOAD54 THAT ENCODES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/955,713
FILING DATE: 23-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/050,124
FILING DATE: 18-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70087
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1594 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-955-713-1

A HUMAN 7-TRANS

```
Query Match      18.1%; Score 188.8; DB 2; Length 1594;
Best Local Similarity 54.1%; Pred. No. 4.9e-45;
Matches 445; Conservative 0; Mismatches 357; Indels 12; Gaps 3;

QY 53 TGGCGCGCTGCTCATTTGGCCCTTTGCTGGCGGCACTAGGCAATGGGTCGCCCTGT 112
DB 519 TGGCACCACATCTGGCCCTGGAGTTTGTCTGGCCCTGGTGGGGAACAGTTTGGCCCTCT 578
QY 113 GTGCTTTCTGCTTCCACATGAAGACCTTGGAGCCAGCAGCTGTTTACCTTTTCAATTTGG 172
DB 579 TCATCTTCTGCATCCACACGCGCCCTGGACCTCCACACAGGTTTCTTGGTACAGCTGG 638
QY 173 CCGTGGCTGATTTCTCTTATGATCTGCTGCTTTTCGACAGACTATTACCTCAGAC 232
DB 639 TGGCGCTGACTTCTCTCTGATCAGCAACCTGCCCTCCGCGTGGACTACTACCTCTCC 698
QY 233 GTAGACACTGGCTTTTGGGACATTCCTTGGCCAGTGGGCGCTTTCAGTTGGCCATGA 292
DB 699 ATGAGACTGGCGCTTTGGGGCTGCTGCTTGCCTGCAAGTCAACCTCTTCTGCTGTGCCACCA 758
QY 293 ACAGGCGCGGAGCATCTGTTCTTACGGTGTGGCTGGGACAGGATTTTCAAGTGG 352
DB 759 ACCGACGCGCAGGTTGTTCTTCTACAGCCATCGCAGCTACCTGAGGTGG 818
QY 353 TCCACCCCGCCACGCGGTGAACACTATCTCCACCCCGGTGGGCTGGCATCTCTGCA 412
DB 819 TGCAGCCCGCCACGCTGCTGAGCGGTCTTCCGTGGGCGAGCTGCCGCGTGGCCGGG 878
QY 413 CCGTGGCGCCCTGGTCTGCTGGAACAGTGTATCTTTTGTGGAGAACCATCTCTGG 472
DB 879 GACTCTGGG---TGGGATCTGCTGCTCTCAACGGGCACTGCTCTGAGCAGCTTCTCCG 935
QY 473 TGAAGAGAGCGCGCTGCTCTGCTGAGAGCTTCAATCATGGAGTGGCCCAATGGCTGGCATG 532
DB 936 GCGCTCTGCTCAGCTACAGGTGGGACGAGCCCTCGGCTCGCTCGCTGGCACC 995
QY 533 ACATCATGTTCCAGCTGAGTCTTTATGCCCCCTCGGCATCATCTTATTTGCTCTTCA 592
DB 996 AGGCACCTGACTCTGCTGAGTCTTCTTCCGCTGACCTGGCGCTCATCTCTTT---GCTATTG 1052
QY 593 AGATTGTTGGAGCTCAGCGGAGGAGCAGCTGGCCAGACAGGCTCGGATGAAGAGG 652
DB 1053 TGAGCATTTGGGCTCACCATCCGGAACCGTGTGCTGGCGGCGAGGCGCCGAGAGG 1112
QY 653 CGACCGGTTTCATCATGCTGCTGCAATTTGTTTCATCATGCTAGCTGCCAGCGCTGT 712
DB 1113 CCATGCTGTGCTGGCCATGGTGGTGGCGCTCTACACCATCTGCTTCTGCGCCAGCATCA 1172
QY 713 CTGCTAGACTCTATTTCTCTGAGAGTGGCCCTGGAGTGGCTTGGATCCCTCTGTCCATG 772
DB 1173 TCTTTGGCATGGCTTCCATGGTGGCTTTCTGGCTGTCCGCTGCCGCTGGACCTCT 1232
QY 773 GGGCCCTGCACATTA-----CCCTCAGCTTACCTACATGAACAGCATGCTGGATCCC 826
DB 1233 GCACAGAGCTCTTCCATGGCTTCCCTGGCCCTTCACTTCAACAGTGTCTCGGACCCG 1292
QY 827 TGGTGTATTTTCAAGCCCTTCTTCCCAA 860
DB 1293 TGCTCTACTGCTTCTAGCCCCCAACTTCTCTCCA 1326

RESULT 5
US-08-955-713-3
: Sequence 3, Application US/08955713
: Patent No. 5955308
: GENERAL INFORMATION:
: APPLICANT: SATHE, GANESH
: APPLICANT: MOONEY, JEFFREY
: APPLICANT: BERGSMAN, DEREK
: APPLICANT: HALSEY, WENDY
: TITLE OF INVENTION: CDNA CLONE HE0AD54 THAT ENCODES A HUMAN 7-TRANS
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
```

```
: ADDRESSEE: RATNER & PRESTIA
: STREET: P.O. BOX 980
: CITY: VALLEY FORGE
: STATE: PA
: COUNTRY: USA
: ZIP: 19482
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/955,713
: FILING DATE: 23-OCT-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/050,124
: FILING DATE: 18-JUN-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: PRESTIA, PAUL F
: REGISTRATION NUMBER: 23,031
: REFERENCE/DOCKET NUMBER: GH-70087
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 610-407-0700
: TELEFAX: 610-407-0701
: TELEX: 846169
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1435 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: US-08-955-713-3
```

Query Match 18.1%; Score 188; DB 2; Length 1435;

Best Local Similarity 54.1%; Pred. No. 7.9e-45;

Matches 440; Conservative 3; Mismatches 359; Indels 12; Gaps 3;

```
QY 53 TGGCGCGCTGCTCATTTGGCCCTTTGCTGGCGGCACTAGGCAATGGGTCGCCCTGT 112
DB 18 TGGCACCACATCTGGCCCTGGAGTTTGTCTGGCCCTGGTGGGGAACAGTTTGGCCCTCT 77
QY 113 GTGCTTTCTGCTTCCACATGAAGACCTTGGAGCCAGCAGCTGTTTACCTTTTCAATTTGG 172
DB 78 TCATCTTCTGCATCCACACGCGCCCTGGACCTCCACACAGGTTCTCTGCTCAGCTGG 137
QY 173 CCGTGGCTGATTTCTCTTATGATCTGCTGCTTTTCGACAGACTATTACCTCAGAC 232
DB 138 TGGCGCTGACTTCTCTCTGATCAGCAACCTGCCCTCCGCTGGACTACTACCTCTCC 197
QY 233 GTAGACACTGGCTTTTGGGACATTCCTTCCGAGTGGGCTCTTCCAGTTGGCCATGA 292
DB 198 ATGAGACTGGCGCTTTGGGGCTGCTGCTGCAAGTCAACCTCTTCTGCTGTGTCNACCA 257
QY 293 ACAGGCGCGGAGCATCTGTTCTTACGTTGGTGGCTGGGACAGGATTTTCAAGTGG 352
DB 258 ACCGCAAGGCGAGGTTGTTCTTCTCAGGCATCGCAGCTCAACCGCTACCTGAAGTGG 317
QY 353 TCCACCCCGCCACGCGGTGAACACTATCTCCACCCCGGTGGCGGTGGCATCTCTGCA 412
DB 318 TGCANCCCGCCACGCTGCTGAAACCGTCTTCCGTGGGGGCACTGCCCGGTGGCGGG 377
QY 413 CCGTGTGGCCCTGGTCTGCTGGAACAGTGTATCTTTTGTGGAGAACCATCTCTGCG 472
DB 378 GAATCTGGG---TGGGATCTCTGCTCTCAACGGGNACTCTGCTCTGAACACTTCTCCG 434
QY 473 TGCAAGAGAGCGCGCTCTCTCTGTGAGAGCTTTCATCATGAGTTCGGCCAAATGGCTG 532
DB 435 GCGCTCTCTGCTCAGCTACAGGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 494
QY 533 ACATCATGTTCCAGCTGGAGTCTTTTATGCGCCCTCGGCATCATCTTATTTTGTCTCTCA 592
```

Db 495 AGGACTGTACCTGCTGGGARTTTTCTGCCACTGGCGCTCATCTCTTTT---GCTATTG 551
QY 593 AGATTCTTTGGAGCTTGAGCGGAGGAGCAGCTGCCAGACAGGCTCGGATGAAGAAG 652
Db 552 TGAGCAATGGGCTCACCATCCGGAACCGTGTCTGGCGGGGAGGAGGCCCGCAGAGGG 611
QY 653 CGACCCGGTTTCATCATGTGGTGGCAATGTGTTCATCATCATGCTACCTGCCAGCGGTGT 712
Db 612 CCATGCGTGTCTGCGCATGGTGGTGGCTGTCTACACCATCTGCTTCTTGGCCAGCATCA 671
QY 713 CTGCTAGACTATATTTCTCTTGACGGTGGCCCTCGAGTGGCTGGATCCCTCTGTCCATG 772
Db 672 TCTTTGGCATGGCTTCATATGGTGGCTTTCTGGCTGTCCGCTGCCGATCCCTGGACCTCT 731
QY 773 GGGCCCTGCACATAA-----CCTCAGCTTACCTACATGAACAGCATGCTGGATCCCG 826
Db 732 GCACACAGCTTCTCATGGCTCCCTGGCTTACCTTACCTCAACAGTGTCTGGACCCCG 791
QY 827 TGGTGTATTATTTTCAAGCCCTCTCTTTCCCAA 860
Db 792 TGTCTACTGCTTCTCTAGCCCCCACTCTCTCCA 825

RESULT 6

US-09-130-749-1

; Sequence 1, Application US/09130749

; Patent No. 6031090

; GENERAL INFORMATION:

; APPLICANT: SHABON, USMAN

; ELISHOURBAGI, NABIL

; TITLE OF INVENTION: MOLECULAR CLONING OF A 7TM

; RECEPTOR (GPR31A)

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: RATNER & PRESTIA

; STREET: P.O. Box 980

; CITY: Valley Forge

; STATE: PA

; COUNTRY: USA

; ZIP: 19482

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA: US/09/130,749

; FILING DATE: 07-Aug-1998

; CLASSIFICATION: UNKNOWN

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: <Unknown>

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: PRESTIA, PAUL F.

; REGISTRATION NUMBER: 23,031

; REFERENCE/DOCKET NUMBER: Gp-70513

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610-407-0700

; TELEFAX: 610-407-0701

; TELEX: 846169

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 960 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-130-749-1

Query Match

Best Local Similarity 13.4%; Score 139.2; DB 3; Length 960;

Matches 427; Conservativity 51.4%; Pred. No. 8.6e-31;

Mismatches 388; Indels 15; Gaps 4;

QY 62 TGCTCATTTGTGGCTTTGTGCTGGCGCACTAGGCAATGGGGTGCCTCTGTGTGGTTTCT 121
Db 59 TGTCTGGGCTGGAGTGTGGGCTGGGCTGCTGGGCAACGGGTGGGCTGTGACCTTCC 118
QY 122 GCTTCCACATGAAGACTGGAAGCCAGCACATGTTTACCTTTTCAATTTGGCGGTGGCTG 181
Db 119 TGTTCGGGTGTCAGGGTGTGGAAGCCGTAGCTGTCTACCTGTCTCAACCTGGCCCTGGCTG 178
QY 182 ATTTCTCTCTATGATCTGCTGCTGCTTTTCGGACACACTATTACCTCAGACGTAGACACT 241
Db 179 ACCTGTGTGGCTGGCTGCTGCTTCTGCTGCTGCTTCTACCTCAGGCTCCAGGCTT 238
QY 242 GGGCTTTTGGGGACATTCCTCGCGAGTGGGGTGTACGTTGGGCATGAACAGGGCCG 301
Db 239 GGCATCTGGCCCTGTGGGCTGCTGGCCCTGCTGCTGCTGGACCTCAGCGCAGCG 298
QY 302 GGAGCATCGTGTTCCTTACGGTGGTGGCTGGCGACAGGTATTTCAAAGTGGTCCACCCCT 361
Db 299 TGGGATGCTGCTTCTTGGCGCGCTGGCTTGGACCGGTACCTCCGTGTGGTCCACCCCTC 358
QY 362 ACCACGGGTGAACACTATCTCCACCGGGTGGCGCTGGCATCGTCTGCACCTGTGGG 421
Db 359 GGCTTAAGGTCAACCTGCTGTCTCTCAGCGCGCTTGGGGGTCTCGGGCTCTGCTGGC 418
QY 422 CCCTGGTCACTCTGGGAACAGTGTATCTTTTGTGGAGAACCACATCTCTGCTGCAAGAGA 481
Db 419 TCCTGATGGTGGCCCTCACCCTGCCCGGGTGTCTCATCTCTGAGGGCGCCAGACTCCA 478
QY 482 CGGCCCTCTCTCTGTGAGAGCTTCATCATGAGTGGG-----CAATGGCTGGCATGACA 535
Db 479 CCAGGTGCCACAGTCTTCTACTCCAGGGCAGACGGCTCCTTCAGCATCATCTGGCAGAAG 538
QY 536 TCATGTTCCAGCTGGAGTTCCTTTATGCCCTCGGCATCATCTTTATTTTGTCTCTCAAGA 595
Db 539 CACTCTCTCTGCTTGTGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 598
QY 596 TTGTTTGGAGCTTGAGCGGAGG---CAGCAGCTGGCGACAGAGGCTCGGATGAAGAAGG 652
Db 599 TCATCAGGCTCTCCAGAAAAGACTCCGGAGGCTGAGAACAAGCCCAAGCTTCAGGGGG 658
QY 653 CGACCCGGTTTCATCATGTGGTGGCAATGTGTTCATCATCATGCTACCTGCCAGCGGTGT 712
Db 659 CCCAGGCACCTGGTGCACCTTGGTGGTGTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCT 718
QY 713 CTGCTAGACT-CTATTTCTCTTGAGCGGTGCCCTCGAGTGGCTGGCATCCTCTGTCCAT 771
Db 719 TGCCAGAGTCTGTGATGCACATCTTCCAGAACTCTGGGAGCTGCAGGGCCCTTTGTGCAG 778
QY 772 GGGGCT-----GCACATAACCTCAGCTTACCTACATGAACAGCATGCTGGATCCCC 826
Db 779 TGCTCATACCTCGGATGTCACGGGAGCTCACCCTACCTGACAGTGTGCTCAACCCCG 838
QY 827 TGGTGTATTATTTTCAAGCCCTCTCTTTTCCCAAAATCTTACAAAGCTC 876
Db 839 TGTATACTGCTTCTCCAGCCCCACCTTTCAGGAGCTCTATCGGAGGCTC 888

RESULT 7

US-09-130-749-1

; Sequence 1, Application US/09130749

; Patent No. 6031344

; GENERAL INFORMATION:

; APPLICANT: SHABON, USMAN

; ELISHOURBAGI, NABIL

; TITLE OF INVENTION: MOLECULAR CLONING OF A 7TM

; RECEPTOR (GPR31A)

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: RATNER & PRESTIA

; STREET: P.O. Box 980

; CITY: Valley Forge

; STATE: PA

QY 48 GGTGATGCCGCCCTGCTCATTTGTCGCTTGTGCTGGCGCACTAGGCAATGGGTGCG 107
Db 796 GCTGTTGGCTCTCTTCTACCTTCTGATTTATCTCTGGCTTTAGTTGGCAATACCTGGC 855
QY 108 CTTGCTGCTGTTCTGCTTCCACATGAAGACCTGGGAAGCCAGCAGCTGTTTACCTTTCAA 167
Db 856 TCTGTGGCTTTTTCATCCGAGACACAAAGTCCGGAGCCCGGCAAGCTGTTCTCTGATGA 915
QY 168 TTTGGCGCTGGCTGATTTCTCTCTTATGATCTGCTTGTGCTGCTTTTTCGGACAGACTATTACCT 227
Db 916 TCTGGCGCTGGCGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 975
QY 228 CAGACCTAGACACTGGGCTTTTGGGACATTCCTCCGAGTGGGCTCTTTCACGTTGGC 287
Db 976 CTTGGGAACCACTGGCCATTTGGGAAATCGCATCCGCTCTCACCGGCTTCTCTCTTA 1035
QY 288 CATGACAGGCGCGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTTCAA 347
Db 1036 CTTCAACATGTACGCCAGCATCTACTTCTCCTCCTGATCGATCGATCGATCGATCGAT 407
QY 348 AGTGTGCTCCACCCACACGCGGTGAACACTATCTCCACCGGGTGGCGCTGGCATCTCT 467
Db 1156 TGCCTTCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1215
QY 468 CTGCGTGCAAGACAGCGCGCTCTCTCTGTGAGAGCTTTCATCATGGAGTTCGGCCAATGCCTG 527
Db 1216 CTTGCGAGACCAACACACAGGTGGTCTGCTGACAGCTGTACCGGGAAGGCTTCCACCA 1275
QY 528 GCATGACATCATGTTCCAGCTGGAGTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 587
Db 1276 TGCCTTCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1335
QY 588 CTTCAAGATTTGTTGAGCTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 647
Db 1336 CTTGCTGATCATCGCGAGCTGCGGAGGCGCTGCGTGGAGAGCGCTCAAGACCAA 1395
QY 648 GAAGCGACCGGTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 698
Db 1396 GGCAGTGCCGATGATGCCATAGTGTGGCCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1455
QY 699 CTTGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 758
Db 1456 CGTCAACCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1515
QY 759 TCCCTGTGCCATGGGCGCTGACATAACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 818
Db 1516 GCGCATCTTGGCGCTGGCAACCGCATCACCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1575
QY 819 GGATCCCTGGTATATTATT 839
Db 1576 CGACCCCATCATGATTTCTT 1596

RESULT 9

US-09-299-843A-43
; Sequence 43, Application US/09299843A
; Patent No. 6107475
; GENERAL INFORMATION:
; APPLICANT: Godiska, Ronald
; APPLICANT: Gray, Patrick W.
; APPLICANT: Schweikart, Vicki L.
; TITLE OF INVENTION: No. 6107475el Seven Transmembrane Receptors
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois

; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM: Floppy disk
; MEDIUM TYPE: IBM PC compatible
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,843A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/088,337
; FILING DATE: 01-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/153,848
; FILING DATE: 17-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/977,452
; FILING DATE: 17-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Jill E. Uhl
; REGISTRATION NUMBER: 43,213
; REFERENCE/DOCKET NUMBER: 27866/32059B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX:
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1901 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 701..1717
; US-09-299-843A-43

Query Match 10.5%; Score 109; DB 3; Length 1901;
Best Local Similarity 47.1%; Pred. No. 6.6e-22;
Matches 377; Conservative 0; Mismatches 415; Indels 9; Gaps 1;
QY 48 GGTGATGCCGCCCTGCTCATTTGTCGCTTGTGCTGGCGCACTAGGCAATGGGTGCG 107
Db 796 GCTGTTGGCTCTCTTCTACCTTCTGATTTATCTCTGGCTTTAGTTGGCAATACCTGGC 855
QY 108 CTTGCTGCTGTTCTGCTTCCACATGAAGACCTGGGAAGCCAGCAGCTGTTTACCTTTCAA 167
Db 856 TCTGTGGCTTTTTCATCCGAGACACAAAGTCCGGAGCCCGGCAAGCTGTTCTCTGATGA 915
QY 168 TTTGGCGCTGGCTGATTTCTCTCTTATGATCTGCTTGTGCTGCTTTTTCGGACAGACTATTACCT 227
Db 916 TCTGTGGCGCTGGCGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 975
QY 228 CAGACCTAGACACTGGGCTTTTGGGACATTCCTCCGAGTGGGCTCTTTCAGTTGGC 287
Db 976 CTTGGGAACCACTGGCCATTTGGGAAATCGCATCCGCTCTCACCGGCTTCTCTCTTA 1035
QY 288 CATGACAGGCGCGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTTCAA 347
Db 1036 CTTCAACATGTACGCCAGCATCTACTTCTCCTGATCGATCGATCGATCGATCGATCGAT 407
QY 348 AGTGTGCTCCACCCACACGCGGTGAACACTATCTCCACCGGGTGGCGCTGGCATCTCT 467
Db 1096 CATTTGTGACCCCGTCAAGTCCCTCAAGCTCCGACGCGCCCTCTACGACACACCTGGCGCTG 1155
QY 408 CTGACACCTGTGGGCGCTTGGTATCTCTGGGAACAGAGTATCTTTTGTGTGAGAACCATCT 467
Db 1156 TGCCTTCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1215
QY 468 CTGCGTGCAAGACAGCGCGCTCTCTCTGTGAGAGCTTTCATCATGGAGTTCGGCCAATGCCTG 527

Db 1216 CGTGCAGACCAACACACGCGTGGTCTGCCTGCAGCTGTACCGGGAGAGGCGCTCCACCA 1275
QY 528 GCATGACATCATGTTCCAGCTGGAGTTCTTTATGCCCTCGGCATCATCTTATTTTGGTC 587
Db 1276 TGGCCTGGTCCCTGGCAGTGGCCTTCACTTCCCTGTTTCATCACCAGGTCACCTGCTA 1335
QY 588 CTCAAGATTGTTGGAGCCTGAGCGGAGCGACAGCTGGCCAGACAGGCTCGGATGAA 647
Db 1336 CCTGCTGATCATCGCAGCCTCGCGCAGGCGCTCGGTGGAGAGCGCCTCAAGACCAA 1395
QY 648 GAAGCGACCGGTTTCATCATGTTGGTGGCAATTGT-----GTTTCATCATGCTA 698
Db 1396 GGCAGTGGCATGATGCCATAGTGTGGCATCTTCTGCTGTGCTTCTGCTGCTTACCA 1455
QY 699 CTTGCCAGGCTGCTCTAGACTTATTTCTCTGACAGCTGCTGCGGCTCGGCTCGGA 758
Db 1456 CGTCAACCGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1515
QY 759 TCCCTCTGTCATCGGCGCTGACATAAGCCTCAGCTTACCTTACATGACAGCATGCT 818
Db 1516 GCGCATCTGCGCCTGCGCAACCGCATCCTCTGCTCACCAGCCTCAACGGGCACT 1575
QY 819 GGATCCCTGGTGTATTATTT 839
Db 1576 CGACCCCATCATGTTATTTCTT 1596

RESULT 10

US-09-088-337B-43
; Sequence 43, Application US/09088337B
; Patent No. 6348574
; GENERAL INFORMATION:
; APPLICANT: Godiska, Ronald
; Gray, Patrick W.
; Schwellkart, Vicki L.
; TITLE OF INVENTION: No. 6348574el Seven Transmembrane Receptors
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; City: Chicago
; State: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/088,337B
; FILING DATE: 01-Jun-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/153,848
; FILING DATE: 17-NOV-1993
; APPLICATION NUMBER: US 07/977,452
; FILING DATE: 17-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6348574and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31794
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1901 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 701..1717
; SEQUENCE DESCRIPTION: SEQ ID NO: 43:
US-09-088-337B-43

Query Match 10.5%; Score 109; DB 4; Length 1901;
Best Local Similarity 47.1%; Pred. No. 6.6e-22;
Matches 377; Conservative 0; Mismatches 415; Indels 9; Gaps 1;

QY 48 GGTGATCGCCCGCTGCTCATGTTGGCTTTGCTGGCGCAGCTAGCAATAGGGTGGC 107
Db 796 GCTGTCGCCCTCTTACCTTCTGGATTCTTATCTTGGCTTTAGTTGGCAATACCTGGC 855
QY 108 CCTGTGTGGTTTCTGCTTCCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTTCAA 167
Db 856 TCTGTGCTTTTTCATCCGAGACCAAGTCGGGACCCCGGCCAACGTGTTCTGTGATGCA 915
QY 168 TTTGGCGTGGCTGATTTCTTCTTATGATCTGCTGCTTTTCGGACAGACTATTACCT 227
Db 916 TCTGCGCTGGCCGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 975
QY 228 CAGACGTAGACACTGGGCTTTTGGGACATTTCCCTGCGGAGTGGGGCTCTTTCAGTTGGC 287
Db 976 CTCTGGGAACCACTGGCCATTTGGGAAATCGCATGCGCTCTCACCGGCTTCTCTTCTA 1035
QY 288 CATGAACAGGCGCGGAGCATCGTTTCTTACGGTGGTGGCTGGGACAGGTATTTCAA 347
Db 1036 CCTCAACATGTACGCCAGCATCTACTTCTTCCATCAGCTGATCAGCGCCGACCGTTTCC 1095
QY 348 AGTGTCTCACCCCCACCCAGCGGTGAACACTATCTCTCACCCGGGTGGCGCTGGCATCGT 407
Db 1096 CATGTGCACCCGGTCAAGTCCCTCAAGCTCCGCGAGGCCCTCTACGCACACACTGGCCTG 1155
QY 408 CTGCACCTCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 467
Db 1156 TGCCTTCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1215
QY 468 CTGCGTCAAGAGACGCGCTCTCTGTGAGAGCTTTCATCATGAGTGGGCAATGGCTG 527
Db 1216 CGTGCAGACCAACACAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1275
QY 528 GCATGACATCATGTTCCAGCTGGAGTTCTTTATGCCCTCGGCATCATCTTATTTTGTCT 587
Db 1276 TGGCCTGGTCTCCCTGGCAGTGGCCTTCACTTCCCTGTTTCATCACCAGGTCACCTGCTA 1335
QY 588 CTTCAAGATTGTTGGAGCCTGAGCGGAGCGACAGCTGGCCAGACAGGCTCGGATGAA 647
Db 1336 CCTGCTGATCATCGCAGCCTGCGGAGGCGCTGCGTGTGGAGAGGCGCTTCAAGACCAA 1395
QY 648 GAAGCGACCGGTTTCATCATGTTGGTGGCAATTGT-----GTTTCATCATGCTA 698
Db 1396 GGCAGTGGCATGATGCCATAGTGTGGCCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1455
QY 699 CCTGCCAGGCTGCTGCTAGACTTATTTCTTGGAGGCTGGCCCTCGAGTGGCTGGCA 758
Db 1456 CGTCAACCGCTCCGCTTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1515
QY 759 TCCCTCTGTCATCGGCGCTGACATAAGCCTCAGCTTACCTTACATGACAGCATGCT 818
Db 1516 GCGCATCTGCGCCTGGCAACCGCATCCTCTGCTCACCAGCCTCAACGGGCACT 1575
QY 819 GGATCCCTGGTGTATTATTT 839
Db 1576 CGACCCCATCATGTTATTTCTT 1596

RESULT 11
PCT-US93-11153-43
; Sequence 43, Application PC/TUS9311153
; GENERAL INFORMATION:

APPLICANT: Godiska, Ronald
 APPLICANT: Gray, Patrick W.
 APPLICANT: Schweikart, Vicki L.
 TITLE OF INVENTION: Novel Seven Transmembrane Receptors
 NUMBER OF SEQUENCES: 64

Query Match	10.5%	Score 109;	DB 5;
Best Local Similarity	47.1%;	Pred. No. 6.6e-22;	
Matches 377;	Conservative	0;	Mismatches 415;
			Indels 9;
			Gaps 1;

Qy	48	GGTGATGCCCGCGCTGCTCATTTGTGCGCCTTTGTGCTGGCGCACTAGGCAATAGGGTCGC	107
Db	796	GC TTGTTCCGCTCCTCTTACTCTGGAATTTATCCTGCTTTAGTTGGCAATACCCGTGC	855
Qy	108	CTGTGTGTTTCTGCTTCCACATCAAGACCTGGAGCCCGAGCACTGTTTACCTTTTCAA	167
Db	856	TC TGTGCGCTTTTTCATCCGAGACCAAGTCCGGGACCCCGGCCAACGTGTTCCTGATGCA	915
Qy	168	TTTGGCGGTGGCTGATTTCTCTTATGATCGCTGCGCTTTTCGAGACAGCACTATTACCT	227
Db	916	TC TGGCGGTGGCGGACTTCTCGTGGGTGCTGGTCTGCCCAACCCGCTGGTCTACCACTT	975
Qy	228	CAGACGTAGACACTGGGCTTTTGGGGAGATTCCTTCGCGAGTGGGGCTCTTCAGTTTGGC	287
Db	976	CTCTGGGAACCACTGGCCATTTTGGGGAATCGCATGCGCTCTACCGGCTTCCTCTCTA	1035
Qy	288	CATGAACAGGGCGGGAGCATCGTTCCTTACGGTGGTGGCTCGGACAGGTATTTTCAA	347
Db	1036	CCTCAACATGTACGCCAGCATCTACTTCTACCTTGCATCAGCGCCAGCCGTTTTCCTGGC	1095
Qy	348	AGTGGTCCACCCGCCACCGGTGAACACTATCTCACCCGGGTGGGGCTGGCATCGT	407
Db	1096	CATTGTGTCACCGGTTCAAGTCCCTCAAGTCCGAGGCCCTCTACGCACACCTGGGCTG	1155

Qy	408	CTGCACCCCTGTGGGCCCTCGTCAATCCTGGGAACAAGTATGATCTTTTGCTGGAGAACAACCATCT	467
Db	1156	TGCCTTCCTGTGGGTGGTGGCTGTGGCCATGGCCCCGCTGTGTGGTAGCCCCACAGAC	1215
Qy	468	CTGGGTGCAAGACAGCGCGCTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTG	527
Db	1216	CGTGCGAACCACACACCGGTGGTCTGCCTGCAGCTGTACCGGGAGAAAGGCTCTCCCACCA	1275
Qy	528	GCATGACATCATGTTCCAGCTGGAGTTCTTTATGCCCCCTCGGCATCATCTTTATTTTGCTC	587
Db	1276	TGCCCTGGTCTCCCTGGCAGTGGSCCTTCACCTTCCTCCGTTTCATCACCGGTACCTGCTA	1335
Qy	588	CTTCAAGATTGTTTGAGGCTGAGCGGAGGAGCAGACTGGCCAGACAGAGCTCGGATGAA	647
Db	1336	CCTGCTCATCATCCGACGCTCGCGGAGGCGCTGCGTGTGGAGAAGCGCCTCAAAGACCA	1395
Qy	648	GAAGGCGACCGGTTTCATCATGGTGTGGCAATTCG-----GTTTCATCACATGCTA	698
Db	1396	GCGATGCGGATCATGCCATAGTGTGGCCATCTTCTGCTGCTTCGTGCGCCTACCA	1455
Qy	699	CCTGCCAGCGTCTCTAGACTCTATTTCCTCTGGACGGTGGCCCTCGAGTGCCTCGCA	758
Db	1456	CGTCAACCGCTCCGCTACGTGTGCACCTACCGCAGCCATGGGCGCTCTCGCGCACCCA	1515
Qy	759	TCCTCTGTCCATGGGGCCCTGCACATACCTCAGCTTCACCTACATGAACACGATGCT	818
Db	1516	GCGCATCCTGGCCCTGGCAACCGCATACCTCCTGCTCCACACGCTCAAACGGGGCACT	1575
Qy	819	GGATCCCGCTGGTATTATT 839	
Db	1576	CGACCCCATCATGATATTCTT 1596	
 RESULT 12 PCT-US95-07180-1 ; Sequence 1, Application PC/TUS9507180 ; GENERAL INFORMATION: ; APPLICANT: LI, YI ; APPLICANT: GOCAYNE, JEANINE D ; APPLICANT: RUBEN, STEVEN M ; TITLE OF INVENTION: G-PROTEIN RECEPTOR HIBEB69 ; NUMBER OF SEQUENCES: 9 ; CORRESPONDENCE ADDRESS: ; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI, ; ADDRESSEE: STEWART & OLSTEIN ; STREET: 6 BECKER FARM ROAD ; CITY: ROSELAND ; STATE: NJ ; COUNTRY: US ; ZIP: 07068 ; COMPUTER READABLE FORM: ; MEDIUM TYPE: Floppy disk ; COMPUTER: IBM PC compatible ; OPERATING SYSTEM: PC-DOS/MS-DOS ; SOFTWARE: Patent In Release #1.0, Version #1.30 ; CURRENT APPLICATION DATA: ; APPLICATION NUMBER: PCT/US95/07180 ; FILING DATE: 06-JUNE-1995 ; CLASSIFICATION: ; ATTORNEY/AGENT INFORMATION: ; NAME: MULLINS, J.G. ; REGISTRATION NUMBER: 30,073 ; REFERENCE/DOCKET NUMBER: 325800-366 ; TELECOMMUNICATION INFORMATION: ; TELEPHONE: 201-994-1700 ; TELEFAX: 201-994-1744 ; INFORMATION FOR SEQ ID NO: 1: ; SEQUENCE CHARACTERISTICS: ; LENGTH: 2453 base pairs ; TYPE: nucleic acid ; STRANDEDNESS: single ; TOPOLOGY: linear ; MOLECULE TYPE: DNA (genomic)			

FEATURE:
NAME/KEY: CDS
LOCATION: 548..1564
PCT-US95-07180-1

Query Match 10.5%; Score 109; DB 5; Length 2453;
Best Local Similarity 47.1%; Pred. No. 7.5e-22;
Matches 377; Conservative 0; Mismatches 415; Indels 9; Gaps 1;

Qy 48 GGTGATGCGCGCGTCTCATTTGGGCTTTGTCTGGGCGCATAGGCAATGGGTCGC 107
Db 643 GCTGTCGCTCCTTACCTCTGGATTATCTGCTGGGCTTTAGTTGGCAATACCCCTGGC 702
Qy 108 CCTGTGGTTCCTGCTTCACATGAGACCTGGAAGCCGACGACTGTTACCTTTTCAA 167
Db 703 TCTGTGCTTTTCATCCGAGACCAAGTCGGGACCCCGGCAACGTGTCTTGATGCA 762
Qy 168 TTGGCGGTGGCTGATTTCTTCCTATGATCTGCTTCCCTTTTCGAGACACTATTACCT 227
Db 763 TCTGGCGTGGCGGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 822
Qy 228 CAGACGTAGACATGGGCTTTTGGGACATTCCTGCGGAGTGGGCTCTTCAGTTGGC 287
Db 823 CTCTGGGAACACATGGGCAATTTGGGAAATCGCATGCCGTCTACCGGCTTCTCTTCTA 882
Qy 288 CATGAACAGGCGCGGAGCATGCTGTTCTTACGGTGGTGGCTCGGACAGGATTTCAA 347
Db 883 CCTCAACATGTAGCCAGCATCTACTTCTTCCCTGATCATGCGCCGACGTTTCTTGGC 942
Qy 348 AGTGGTCCACCCCGGAGGAGTATCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 407
Db 943 CATTTGTGACCGGCTCAAGTCCCTTACAGCTCCGCGAGGCGCTCTACGACACCTGGCGTG 1002
Qy 408 CTGACACCTGTGGGCGCTGTCATCTCTGGGACAGTATCTTTTGTGGAGACCATCT 467
Db 1003 TGGCTTCTGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1062
Qy 468 CTGCGTCAAGAGACGCGCTCTCTGTGAGAGCTTCATCATGGAGTGGGCAATGGCTG 527
Db 1063 CGTGCACCAACACACAGCGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1122
Qy 528 GCATGACATCATGTTCCAGTGGAGTCTTTATGCGCCCTCGGCAATCATTTATTTGCTC 587
Db 1123 TGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1182
Qy 588 CTTCAGATGTTTGGAGCTGAGCGGAGGAGCAGCAGCTGSCCAGACAGGCTCGATGAA 647
Db 1183 CCTGCTGATCATCCGACGCTGCGGCGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1242
Qy 648 GAAGCGACCGGCTTCATCATGTTGGTGGCAATTGT-----GTTTCATCATGCTA 698
Db 1243 GGCAGTGGCATGATGCCATAGTGTGGCCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1302
Qy 699 CCGCCAGCGGTGCTGCTAGACTCTATTTCCTGTCGACGCGTGGCCCTCGAGTGGCTCGGA 758
Db 1303 CGTCAACCGTCCGCTCTACGCTGTGCACTACCGGAGCATGGGCGCTCTCGGCGCACCA 1362
Qy 759 TCCCTCTGTCATGGGCGCTGCAATACCTCAGCTTACCTACATGACACATGCT 818
Db 1363 GCGCATCTGCGCCCTGGAACCGCATCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1422
Qy 819 GGATCCCGCTGCTGATTTATT 839
Db 1423 CGACCCCATCATGATTTCTT 1443

RESULT 13
US-08-724-974A-1
Sequence 1, Application US/08724974A
Patent No. 5912335
GENERAL INFORMATION:
APPLICANT: Derk J. Bergsma, Catherine E. Ellis
TITLE OF INVENTION: A No. 5912335el G-Protein Coupled Receptor

TITLE OF INVENTION: HUVCT36
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road, P.O. Box 1539
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
COMPUTER: IBM 486
OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,974A
FILING DATE: October 3, 1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: William T. Han
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: ATG50022
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610 270 5024
TELEFAX: 610 270 5090
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1597
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
ANTI-SENSE: No
US-08-724-974A-1

Query Match 10.0%; Score 104.2; DB 2; Length 1597;
Best Local Similarity 46.9%; Pred. No. 1.5e-20;
Matches 399; Conservative 0; Mismatches 443; Indels 9; Gaps 2;
Qy 5 ACAACGGGTGCTGCTGCGCATGAGGGGACACCATCTCCAGGTGATGCGCGCTGC 64
Db 361 ACAACTCCTCGATGAGCTGTACCATCGACCATACCATCCACGACGCTGGCCGGTG 420
Qy 65 TCATTGTGGCC---TTTGTGCTGGCGCAGTACGCAATGGGTGCGCCCTGTGTGTTCT 121
Db 421 TCTATGTTACCGTGTGCTGGGCTTCCCGGCAACTGCTGTCTCTTACTTTCGGCT 480
Qy 122 GCTTCCACATGAAGACCTGGAAGCCAGACACTGTGTTTACCTTTTCAATTTGGCCGTGCTG 181
Db 481 ACCTGCAGATCAAGGCCCGGACGAGCTGGCGGTGTACCTGTGCAACTGACGCTGGCCG 540
Qy 182 ATTTCTCTCTTATGATGCTGCTGCTTTTCGGACAGACTATTACCTCAGACGTAGACACT 241
Db 541 ACCTCTTCTACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Qy 242 GGGCTTTTGGGACATTCCTGCTGGGCTGCTTTCAGCTTGGCCATGACAGGGCGC 301
Db 601 GGTCTCAGCGGACCTGCTGCTGCGAGGTGTGGGCACTCTCTCTGTAGGAGAACTATACA 660
Qy 302 GGAGCATGCTGTTCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361
Db 661 TCAGCGTGGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Qy 362 ACCACGCGGTGAACACTATCTCCACCGGGTGGCGGTGGCGGTGCTGCTGCTGCTGCTGCTGCTGCT 421
Db 721 TCGGCTTCCACAGTTCGCGACCTGGAAGCGCGCGCTGCGGCTGAGCGGTGCTGCTGCTGCTGCTGCT 780
Qy 422 CCCTGGTCTCCTGGGAAACAGTGTATCTTTTGTGGAGAACCATCTCTGCTGCTGCTGCTGCTGCTGCT 481
Db 781 CCAAGGAGCTGCTGACCAAGCATCTACTTCTGATGTCGACGAGGAGGTCTATCGAGGACGAGA 840

```
QY 482 -----CGGCCGCTCTCTGTGAGAGCTTCAATCATGGAGTGGCCAAATGGCTGCATGACA 535
Db 841 ACCAGCACCGCGTGTCTTTGAGCACTACCCCATCCAGGCAATGCGAGCGGCCCACTCAACT 900
QY 536 TCATGTTCCAGCTGAGTCTTTATGCCCCCTCGGCATCATCTTATTTTGTCTCCTTCAAGA 595
Db 901 ACTACCGCTTCCCTGGTGGGCTTCTCTTCCCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
QY 596 TTGTTTGGAGCTGAGCGGAGGCGAGCTGCGCCAGACAGCTCGGATCAAGAAGGCGGA 655
Db 961 GCATCTGCGCGCGTGGCGGAGCCAGCCAGCCAGAGAGCGCGCAAGACCATGATCC 1020
QY 656 CCGCGTTATCATGGTGGTGGCAATTTGTTCATCATCATCTACCTACCTGCCAGCGTGTCTG 715
Db 1021 AGCGGCTGCTGCTGAGCAGCGTGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
QY 716 CTAGACTCTATTTCCTGTGACGGTGGCTGAGTGGCTGAGTGGCTGAGTGGCTGAGTGGG 775
Db 1081 TGTGCTGTGGCGAGCGTCTGGGAGCGGAGCTGCGGACTTTCGCAAGGCGGTTTCAACG 1140
QY 776 CCCTGCACATAACCCCTCAGCTTACCTTACATGAACAGCATGCTGGATCCCTGTGTATT 835
Db 1141 CCTACCACTTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1200
QY 836 ATTTTCAAGC 846
Db 1201 GCTTGGTCAAGC 1211
```

RESULT 14

```
US-08-442-134A-1
; Sequence 1, Application US/08442134A
; Patent No. 5596088
; GENERAL INFORMATION:
; APPLICANT: Boucher, Richard C.
; APPLICANT: Weisman, Gary A.
; APPLICANT: Turner, John T.
; APPLICANT: Harden, Thomas K.
; APPLICANT: Parr, Claude E.
; APPLICANT: Sullivan, Daniel M.
; APPLICANT: Erb, Laura
; APPLICANT: Lustig, Kevin D.
; TITLE OF INVENTION: DNA Encoding the Human P2U Receptor and
; TITLE OF INVENTION: Null Cells Expressing P2U Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell, Seltzer, Park & Gibson
; STREET: Post Office Drawer 34009
; CITY: Charlotte
; STATE: No. 5596088th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/442,134A
; FILING DATE: 16-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-71A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-420-2200
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1842 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 57..1181
; US-08-442-134A-1

Query Match      8.9%; Score 92.6; DB 1; Length 1842;
Best Local Similarity 51.3%; Pred. No. 3.6e-17;
Matches 215; Conservative 0; Mismatches 204; Indels 0; Gaps 0;

QY 10 GGGTGTGCTGTCGCGCATCGAGGGGAGACCATCTCCAGGTGATGCGCGCGTGTCTATT 69
Db 120 GGCTACAGGTGCGGCTTCAACGAGGACTTCAAGTACGTGCTGCTGCTGCTGCTGCTGCTG 179
QY 70 GTGGCTTTTGTGCTGGCGCAGCTAGGCAANTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 129
Db 180 GTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 239
QY 130 ATGAAGACCTGGAAGCCAGCACTGTTTACCTTTTCAATTTGGCCGCTGGCTGATTTCTCT 189
Db 240 CTCAAGACCTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 299
QY 190 CTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 249
Db 300 TATGCGGCTTCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 359
QY 250 GGGGACATTCCTGCGAGTGGGCTCTTACAGTTGGCCATGAACAGGGCGGGAGCATC 309
Db 360 AGCAGCGTGTCTGCAAGTGGTGGCTTCTCTTCTTACACCACTTTTACTGCGAGCATC 419
QY 310 GTGTTCTTACGGTGGTGGCTGCGGAGCAGGTATTTCAAGTGGTGTCCACCCACACGCG 369
Db 420 CTCTTCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 479
QY 370 GTGAACACTATCTCCACCGGGTGGGCTGCGATCGTGTGCAACCTGTGGGCGCTTGGT 428
Db 480 CTGCGTGGGCGGCGCGCTAGCGCTGCGCGGGTGGCGGGCGGCGGTGGTGGTGGTGGTGGT 538

RESULT 15
US-08-444-581B-1
; Sequence 1, Application US/08444581B
; Patent No. 5607836
; GENERAL INFORMATION:
; APPLICANT: Boucher, Richard C.
; APPLICANT: Weisman, Gary A.
; APPLICANT: Turner, John T.
; APPLICANT: Harden, Thomas K.
; APPLICANT: Parr, Claude E.
; APPLICANT: Sullivan, Daniel M.
; APPLICANT: Erb, Laura
; APPLICANT: Lustig, Kevin D.
; TITLE OF INVENTION: DNA Encoding the Human P2U Receptor and
; TITLE OF INVENTION: Null Cells Expressing P2U Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell, Seltzer, Park & Gibson
; STREET: Post Office Drawer 34009
; CITY: Charlotte
; STATE: No. 5607836th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,581B
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: US 08/442,134
; FILING DATE: 16-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-71A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-420-2200
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1842 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 57..1181
US-08-444-581B-1

Query Match      8.9%; Score 92.6; DB 1; Length 1842;
Best Local Similarity 51.3%; Pred. No. 3.6e-17;
Matches 215; Conservative 0; Mismatches 204; Indels 0; Gaps 0;

Qy 10 GGGTCGTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATGCCGCCGCTGCTCAT 69
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 120 GGCTACAGGTGCCGCTTCAACGAGGACTTCAAGTACGTGCTGCTGCCTGTCTACGGC 179

Qy 70 GTGGCCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTGTGTTTCTGCTTCCAC 129
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 180 GTGGTGTGCTGCTGGGCTGTGCTGAACGCCGGTGGGCTCTACATCTTCTTGTGGCGC 239

Qy 130 ATGAAGACCTGGAAGCCCGACGACTGTTTACCTTTTCAATTTGGCCGTGGCTGATTCCTC 189
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 240 CTCAGACCTGGAATGCGTCCACCACATATATGTTCCACCTGGCTGTGCTGATGCACTG 299

Qy 190 CTTATGATCTGCTGCCCTTTTCGGACAGACTATTACCTCAGACGTFAGACACTGGGCTTTT 249
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 300 TATGGCGCCTCCCTGCCGCTGCTGTCTATTACTAGCCCGCGGGGACCACTGGCCCTTC 359

Qy 250 GGGGACATTCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCGGGAGCATC 309
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 360 AGCAGGTGCTCTGCAAGCTGTGGCTTCTCTTACACCAACCTTTACTGCAGCATC 419

Qy 310 GTGTTCCCTACGGTGGTGGCTGGGACAGGTATTTCAAAGTGGTCCACCCACCACGCG 369
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 420 CTCCTTCCCTACCTGATCAGCGTGCACCGGTCTGTGGCGCTTTACGACCTCTGCGCTCC 479

Qy 370 GTGAACACTATCTCACCCGGGTGGCGCTGGCATCGTCTGCACCCCTGTGGCCCTGGT 428
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 480 CTGGCCTGGGGCGGGCCGCTACGCTCGCGGGGTGGCGGGCGGTGGGTGTTGGT 538
```

Search completed: October 29, 2002, 02:43:00
Job time : 71 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2002, 03:50:26 ; Search time 69 seconds

(without alignments)
867.482 Million cell updates/sec

Title: US-09-886-041-2

Perfect score: 1853

Sequence: 1 MYNSCCRIEGTISQVMP.....ANSFQSDGQWDPHIVEWH 346

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_19:*

- 1: sp.archaea:*
- 2: sp.bacteria:*
- 3: sp.fungi:*
- 4: sp.human:*
- 5: sp.invertebrate:*
- 6: sp.mammal:*
- 7: sp.mhc:*
- 8: sp.organelle:*
- 9: sp.phage:*
- 10: sp.plant:*
- 11: sp.protist:*
- 12: sp.virus:*
- 13: sp.vertebrate:*
- 14: sp.unclassified:*
- 15: sp.rvirus:*
- 16: sp.bacteriap:*
- 17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1853	100.0	346	4 Q9BXC0	Q9bxc0 homo sapien
2	868.5	46.9	360	11 Q9EP66	Q9ep66 mus musculu
3	450.5	24.3	319	4 Q9NQ20	Q9nq20 homo sapien
4	446.5	24.1	319	11 Q9JLS1	Q9jls1 mus musculu
5	371.5	20.0	309	11 Q920A1	Q920a1 mus musculu
6	366.5	19.8	330	4 Q9HCQ2	Q9hcq2 homo sapien
7	366.5	19.8	346	4 Q9NS75	Q9ns75 homo sapien
8	362	19.2	345	6 Q9SNO3	Q9sno3 sus scrofa
9	355.5	19.2	309	11 Q924T9	Q924t9 rattus norv
10	342.5	18.5	357	13 Q9DE05	Q9de05 raja erinac
11	341	18.4	361	13 Q9DX57	Q9dx57 xenopus lae
12	336	18.1	367	4 Q9UE21	Q9ue21 homo sapien
13	334	18.0	288	4 Q9UDZ6	Q9udz6 homo sapien
14	329.5	17.8	380	13 Q9DGQ6	Q9dgq6 carassius a
15	317	17.1	373	11 Q9CPZ4	Q9cpz4 mus musculu
16	310	16.7	377	4 Q96EM8	Q96em8 homo sapien

17	308	16.6	340	6 Q95N02	Q95n02 sus scrofa
18	305	16.5	337	4 Q9Y271	Q9y271 homo sapien
19	305	16.5	372	4 Q9HIC0	Q9hlc0 homo sapien
20	303.5	16.4	359	6 Q9N0U1	Q9n0u1 ovis aries
21	302.5	16.3	346	4 Q96GE0	Q96ge0 homo sapien
22	302.5	16.3	356	4 Q96TF2	Q96tf2 homo sapien
23	300.5	16.2	346	6 Q95KS6	Q95ks6 ovis aries
24	299.5	16.2	346	11 Q91V73	Q91v73 mus musculu
25	296.5	16.0	385	4 Q76067	Q76067 homo sapien
26	296.5	16.0	385	4 Q96RI0	Q96ri0 homo sapien
27	293	15.8	315	6 Q9GRP7	Q9grp7 sus scrofa
28	293	15.8	374	13 Q57466	Q57466 meleagris g
29	293	15.8	385	11 Q9JK40	Q9jk40 mus musculu
30	292.5	15.8	361	6 Q46685	Q46685 bos taurus
31	292	15.8	317	11 Q99MT6	Q99mt6 mus musculu
32	292	15.8	359	6 Q9GLN9	Q9gl9 pan troglod
33	291.5	15.7	358	4 Q96JZ8	Q96jz8 homo sapien
34	291	15.7	359	11 Q9EPP3	Q9eep3 cavia porce
35	290.5	15.7	337	4 Q96P68	Q96p68 homo sapien
36	290.5	15.7	359	11 Q99MT7	Q99mt7 mus musculu
37	289	15.6	343	11 Q9QW32	Q9qw32 rattus sp.
38	289	15.6	352	11 Q9JK47	Q9jk47 mus musculu
39	288.5	15.6	359	11 Q9EQR9	Q9eq9 meriones un
40	288	15.5	355	11 Q9JLY8	Q9jly8 rattus norv
41	287.5	15.5	358	4 Q9BY21	Q9by21 homo sapien
42	286.5	15.5	339	11 Q924T8	Q924t8 rattus norv
43	286.5	15.5	355	4 Q9BYX5	Q9byx5 homo sapien
44	286.5	15.5	423	5 Q964D4	Q964d4 periplaneta
45	285	15.4	339	11 Q9JU71	Q9jj71 mus musculu

ALIGNMENTS

RESULT 1

ID	Q9BXC0	PRELIMINARY;	PRT;	346 AA.
AC	Q9BXC0;			
DT	01-JUN-2001 (Tremblrel. 17, Created)			
DT	01-JUN-2001 (Tremblrel. 17, Last sequence update)			
DT	01-DEC-2001 (Tremblrel. 19, Last annotation update)			
DE	PUTATIVE CHEMOKINE RECEPTOR (G PROTEIN-COUPLED RECEPTOR).			
GN	FKSG80 OR GPR81.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Wang Y.-g, Gong L.;			
RT	"Molecular cloning of FKSG80, a novel gene encoding a putative			
RT	chemokine receptor.;"			
RL	Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21458557; PubMed=11574155;			
RA	Lee D.K., Nguyen T., Lynch K.R., Cheng R., Vanti W.B., Arkhitko O.,			
RA	Lewis T., Evans J.F., George S.R., O'Dowd B.F.;			
RT	"Discovery and mapping of ten novel G protein-coupled receptor			
RT	genes.;"			
RL	Gene 275:83-91(2001).			
DR	EMBL; AF345568; AAK29071.1; -.			
DR	HSSP; P34996; IDDD.			
DR	InterPro: IPR000276; GPCR_Rhodpsn.			
DR	Pfam; PF00001; 7tm.1; 1.			
DR	PRINTS; PR00237; GPCRHHODOPSN.			
DR	PROSITE; PS00237; G_PROTEIN_RECP_FL_1; UNKNOWN_1.			
DR	PROSITE; PS0262; G_PROTEIN_RECP_FL_2; 1.			
KW	Receptor.			
SQ	SEQUENCE 346 AA; 39295 MW; E0DB114EEB3A47A5 CRC64;			

Query Match 100.0%; Score 1853; DB 4; Length 346;

```
Best Local Similarity 100.0%; Pred. No. 2.2e-177;
Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYNGSCCRLEGDTISQVMPPLLIIVAFVIGALGNVALGCGCFHMKTKPSTVYLFENLAVA 60
Db 1 MYNGSCCRLEGDTISQVMPPLLIIVAFVIGALGNVALGCGCFHMKTKPSTVYLFENLAVA 60

Qy 61 DFLLMICLPFRDYYLRRHWAFAFGDIPCRVGLFTLAMNAGSIVELTVVAADRYKVVHP 120
Db 61 DFLLMICLPFRDYYLRRHWAFAFGDIPCRVGLFTLAMNAGSIVELTVVAADRYKVVHP 120

Qy 121 HAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANGWHDIM 180
Db 121 HAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANGWHDIM 180

Qy 181 FOLEFMPGLIIFCSEKIVWSLRRQQLARQARKKATRFIMVVAIVFITCYLPSVSAR 240
Db 181 FOLEFMPGLIIFCSEKIVWSLRRQQLARQARKKATRFIMVVAIVFITCYLPSVSAR 240

Qy 241 LYFLTWVPSSACDPSVHGALHITLSFTYMNMSMLDPLVYFSSPPKFNKIKICS LKPK 300
Db 241 LYFLTWVPSSACDPSVHGALHITLSFTYMNMSMLDPLVYFSSPPKFNKIKICS LKPK 300

Qy 301 QPGHKTQRPPEMPTISNIGRSCISVANSFQSDGQNDPHIVEWH 346
Db 301 QPGHKTQRPPEMPTISNIGRSCISVANSFQSDGQNDPHIVEWH 346

RESULT 2
Q9EP66 PRELIMINARY; PRT; 360 AA.
ID Q9EP66
AC Q9EP66;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE PUTATIVE SEVEN TRANSMEMBRANE SPANNING RECEPTOR.
GN PUMAG OR PUMA-G.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ, AND C57BL/6;
RA Schaub A., Futterer A., Pfeiffer K.;
RT "PUMA-G, an interferon-gamma inducible gene in macrophages is a novel
RT member of the seven transmembrane spanning superfamily.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AJ300199; CAC17791.1; -
DR EMBL; AJ300198; CAC17790.1; -
DR HSP; P34996; 1DD.
DR MGD; MGI:193383; Pumag.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR F1.1; 1.
DR PROSITE; PS0262; G-PROTEIN RECEPTOR F1.2; 1.
DR G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 360 AA; 41400 MW; CCCE52A247577FC CRC64;

Query Match 46.9%; Score 868.5; DB 11; Length 360;
Best Local Similarity 55.6%; Pred. No. 9.3e-79;
Matches 178; Conservative 39; Mismatches 96; Indels 7; Gaps 4;

Qy 3 NG-SCCRLEGDTISQVMPPLLIIVAFVIGALGNVALGCGCFHMKTKPSTVYLFENLAVAD 61
Db 11 NGKNCCVFRDENIAKVLPPVGLGVFLGGLGALWIFCFHLKSKSSKSRIFLFLNLA 70

Qy 62 FLLMCLPFRDYYLRRHWAFAFGDIPCRVGLFTLAMNAGSIVELTVVAADRYKVVHP 121
Db 71 FLLIICLPFLTDYVHNDWRFGGIPCRVGLFTLAMNAGSIVELTVVAADRYKVVHP 130
```

```
Qy 122 HAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANGWHDIMF 181
Db 131 HFLANKISNRTRAAIISCLPGLWLTIGLTVHLLVTNMKTNGEAYVLCSSFSICYNRHWDAMF 190

Qy 182 QLEFMPGLIIFCSEKIVWSLRRQQLARQARKKATRFIMVVAIVFITCYLPSVSARL 241
Db 191 LLEFPLAIILFCSGRIIWSLRQ-OMDRHAKIKRAINFIMVVAIVFITCYLPSVAVRI 249

Qy 242 YFLTWVPS--SACD--PSVHGALHITLSFTYMNMSMLDPLVYFSSPPKFNKIKICS 296
Db 250 RIFWLYKYVNRNCDIYSSVDLAFFTLTSFTYMNMSMLDPLVYFSSPPKFNKICINRC 309

Qy 297 LKPKQPGHKTQRPPEMPTIS 316
Db 310 LRKKTLPEDNNRSTSVELT 329

RESULT 3
Q9NQ20 PRELIMINARY; PRT; 319 AA.
ID Q9NQ20
AC Q9NQ20;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE BA517H2.2 (G PROTEIN-COUPLED RECEPTOR 31).
GN GPR31.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Almeida J.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL121935; CAB99329.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR F1.1; UNKNOWN_1.
DR PROSITE; PS0262; G-PROTEIN RECEPTOR F1.2; 1.
KW Receptor.
SQ SEQUENCE 319 AA; 35074 MW; 2ACD0350AD7FB53A CRC64;

Query Match 24.3%; Score 450.5; DB 4; Length 319;
Best Local Similarity 34.4%; Pred. No. 6e-37;
Matches 101; Conservative 61; Mismatches 121; Indels 11; Gaps 5;

Qy 7 CRIEGDTISQVMPPLLIIVAFVIGALGNVALGCGCFHMKTKPSTVYLFENLAVADFLMI 66
Db 6 CSAPSTVATAVGVLLGLECGLLGNALWLTFLFRVRVWKPYAVYLLNALADLLAA 65

Qy 67 CLPFRDYYLRRHWAFAFGDIPCRVGLFTLAMNAGSIVELTVVAADRYKVVHPHAYNT 126
Db 66 CLPFLAAYFLSLQAWHLGRVGCWALHFLDLDSRSVGMFLAVALDRYLVVHPRLKYNL 125

Qy 127 ISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANG----WHDIMF 181
Db 126 LSPQAALGVSGLVLLMWALTCPLLISE--AAQNSTRCHSF-YSRADGSSIIQEAALS 182

Qy 182 QLEFMPGLIIFCSEKIVWSLRRR-QQLARQARKKATRFIMVVAIVFITCYLPSVSAR 240
Db 183 CLQFVLPGLIVFCNAGIIRALQRLREPKQPKLQRAQALVTLVVVLFALCFPCFLAR 242

Qy 241 --LFLTWVPSACDPSVHGALHITLSFTYMNMSMLDPLVYFSSPPKFNK 292
Db 243 VLMHIFQNLGSCALCAVAHTSDVTGSLTYLHSLVLPVYVFCSSPTFRSSVRRV 296

RESULT 4
Q9JLS1 PRELIMINARY; PRT; 319 AA.
ID Q9JLS1
AC Q9JLS1;
```



```
QY 177 HDIMFQLEFFMPLGLIILFCSEFKIVNSLRQ--QLARQARKKATRFIMVAIVFITCYL 234
Db 186 NYIALVWGCLLPFFTLSCYLLIIRVLLKVEPESGLRSHRKALTTIITLIIFLCLF 245
QY 235 PVSARLYFLMTVPSSACDPVHGALHITLSFTYNSMLDPLVYFSSPKFKYNKIKI 294
Db 246 PYHTLRTVHLTWTWKVGLCKDRHLKALVITLALAAANACFNPLLYFAGENFK--DRLK- 301
QY 295 CSLKPKQPGHKTQ 308
Db 302 SALRKGHPQAKTK 315

RESULT 7
Q9NS75 PRELIMINARY; PRT; 346 AA.
AC Q9NS75;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CYSTEINYL LEUKOTRIENE CYSLT2 RECEPTOR (BA108P5.1) (CYSTEINYL
DE LEUKOTRIENE RECEPTOR TYPE 2).
GN BA108P5.1 OR CYSLT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=20374466; PubMed=10913337;
RA Takasaki J., Kanohara M., Matsumoto M., Saito T., Sugimoto T., Ota T.,
RA Nishikawa T., Kawai Y., Masuho Y., Isogai T., Suzuki Y., Sugano S.,
RA Furuichi K.;
RT "The molecular characterization and tissue distribution of the human
RT cysteinyl leukotriene CYSLT2 receptor.";
RL Biochem. Biophys. Res. Commun. 274:316-322(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX PubMed=10851239;
RA Helse C.E., O'Dowd B.F., Figueroa D.J., Sawyer N., Nguyen T., Im D.S.,
RA Stocco R., Bellefeuille J.N., Abramovita M., Cheng R.,
RA Williams D.L. Jr., Zeng Z., Liu Q., Ma L., Clements M.K., Coulombe N.,
RA Liu Y., Austin C.P., George S.R., O'Neill G.P., Metters K.M.,
RA Lynch K.R., Evans J.F.;
RT "Characterization of the human cysteinyl leukotriene 2 receptor.";
RL J. Biol. Chem. 275:30531-30536(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Dunn M.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Nothacker H.-P., Wang Z., Zhu Y., Civelli O.;
RT "Identification, Molecular Cloning and Characterization of a Second
RT Human Cysteinyl Leukotriene Receptor.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038269; BAB03601.1; -
DR EMBL; AF254664; AAG17281.1; -
DR EMBL; AL137118; CAC29102.1; -
DR EMBL; AF279611; AAK69485.1; -
DR InterPro; IPR004071; Cysleuk_receptor.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR01533; CYSLTRECPTR.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS0262; G_PROTEIN_RECPEP_F1_2; 1.
SQ SEQUENCE 346 AA; 39635 MW; EB54A4A2DDCE5EE4 CRC64;

Query Match 19.8%; Score 366.5; DB 4; Length 346;
Best Local Similarity 29.9%; Pred. No. 1.7e-28;
Matches 94; Conservative 69; Mismatches 134; Indels 17; Gaps 8;
```

```
QY 3 NSCCRIEDDTISQWMPPLLIIVAFVLGALNGVALCGCFHMKTKPST---VYLFNLAIV 59
Db 27 NSRNTIE-NKREFFPIVLIIFWGLNGLSIYVF---LQPKKSTSVNFMNLAI 82
QY 60 ADFLLMICLPRTDYLLRRHWAFGDIPCRVGLFTLANNRAGSIVFLVVAADRIFKVVH 119
Db 83 SDLLFISFLPRADYLLRGSNMIFGDLACRIMSYSLYNNMYSSYFLTVLSVVRFLAMVH 142
QY 120 PHAVNTISTVAAGIVCTLWALVILGVYLLLENHLCVQETAVSC---ESFIMESANGW 176
Db 143 PFRLUHTSIR-SAWILCGIILIMASSIMLLDSSQSGSVTSCLELNLYKIAKLQTM 201
QY 177 HDIMFQLEFFMPLGLIILFCSEFKIVNSLRQ--QLARQARKKATRFIMVAIVFITCYL 234
Db 202 NYIALVWGCLLPFFTLSCYLLIIRVLLKVEPESGLRSHRKALTTIITLIIFLCLF 261
QY 235 PVSARLYFLMTVPSSACDPVHGALHITLSFTYNSMLDPLVYFSSPKFKYNKIKI 294
Db 262 PYHTLRTVHLTWTWKVGLCKDRHLKALVITLALAAANACFNPLLYFAGENFK--DRLK- 317
QY 295 CSLKPKQPGHKTQ 308
Db 318 SALRKGHPQAKTK 331

RESULT 8
Q95N03 PRELIMINARY; PRT; 345 AA.
AC Q95N03;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CYSLT2.
GN CYSLT2.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Kanohara M., Takasaki J., Matsumoto M., Matsumoto S., Saito T.,
RA Ohishi T., Soga T., Matsushime H., Furuichi K.;
RT "Characterization of the cloned rat and porcine cysteinyl leukotriene
RT receptors.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB052662; BAB60817.1; -
SQ SEQUENCE 345 AA; 39410 MW; 5D1B1FB89B95905 CRC64;

Query Match 19.5%; Score 362; DB 6; Length 345;
Best Local Similarity 31.3%; Pred. No. 4.8e-28;
Matches 86; Conservative 59; Mismatches 118; Indels 12; Gaps 5;
```

```
RESULT 9
Q924T9 PRELIMINARY; PRT; 309 AA.
ID Q924T9
AC Q924T9
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CYSLT2.
GN CYSLT2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN
RP SEQUENCE FROM N.A.
RA Kenohara M., Takasaki J., Matsumoto M., Matsumoto S., Saito T.,
RA Ohishi T., Soga T., Matsushime H., Furuchi K.;
RT "Characterization of the cloned rat and porcine cysteine leukotriene
RT receptors."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB052661; BAB60816.1; --
SQ SEQUENCE 309 AA; 35507 MW; A122AC8177879D56 CRC64;

Query Match 19.2%; Score 355.5; DB 11; Length 309;
Best Local Similarity 31.2%; Pred. No. 1.9e-27;
Matches 91; Conservative 57; Mismatches 131; Indels 13; Gaps 7;

Qy 2 YNGSCRTGDTISQVMPPLIVAFVGLGNGVLCGFCFHMKTWKST---VYLFNLA 58
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 10 YSKNCTIE-NFKRDFYPIIILIVFVGLGNGVLCGFCFHMKTWKST---VYLFNLA 65
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 59 VADFLMLICLPFRDYLRRLRHAFGDIQCRVGLFTLAMNRAGSIVFLTVVAADRYFKV 118
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 66 ISDFELFISTLPRADYNFRGSDWIFGDWACRMSVLYNNMTSYIFLTLSIVRFLATA 125
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 119 HPHAVNTISTRVAGIVC-TLWALVILGTIVLLENHLCVQETAVSCSFIMESAN--G 175
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 126 HPFQMLHTSVR-SAWILCGIIVFIMASSGLLKHGQKNNITLCFELNLQRFKNLVI 184
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 176 WHDINFOLEFPNPGIILFCFKIVWSLRQ--QLARQARKKATRIMVVAIVFTCY 233
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 185 LNYIALGVGLFPFLIFTCVLLIIRVLLKVEIPESGPRDQRKALTIVIAMIFLLCF 244
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 234 LPSVSARLYELTWTPSSACDPSVHGALHITLFTYMNMLDPLVYFSSPSF 285
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 245 LPYHALRTHLVTDWADSCMDLHKATVITUTLAAANSCFNPFLYFAGENF 296
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
Q9DE05 PRELIMINARY; PRT; 357 AA.
ID Q9DE05
AC Q9DE05
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE P2Y RECEPTOR.
OS Raja erinacea (Little skate).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalae; Hypnosqualea; Pristiogalea; Batoidae;
OC Rajiformes; Rajidae; Raja.
OX NCBI_TaxID=7782;
RN
RP SEQUENCE FROM N.A.
RA Ballatori N., Boyer J.L., Nathanson M.H.;
RA Dranoff J.A., O'Neill A.F., Franco A.M., Cai S.Y., Connolly G.C.,
RT "A primitive ATP receptor from the little skate Raja erinacea."
RL J. Biol. Chem. 275:30701-30706(2000).
DR EMBL: AF42850; AAG42684.1; --
DR HSSP: P34996; 1DDO.
DR InterPro: IPR000276; GPCR_Rhodpsn.

Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PRO0237; GPCRHHODOPS.
DR PROSITE: PS00237; G_PROTEIN_RECIP_F1_1; UNKNOWN_1.
DR PROSITE: PS0262; G_PROTEIN_RECIP_F1_2; 1.
KW RECEPTOR.
SQ SEQUENCE 357 AA; 41239 MW; 14604EE15DCBDB41 CRC64;

Query Match 18.5%; Score 342.5; DB 13; Length 357;
Best Local Similarity 26.4%; Pred. No. 4.5e-26;
Matches 87; Conservative 69; Mismatches 144; Indels 29; Gaps 9;

Qy 18 MPPLIIVAFVGLGNGVLCGFCFHMKTWKSTVYLFNLAADFLMLICLPFRDYLR 77
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 38 LPIMVIVFTVGTGNSVALWNIHFMRPWSISITIMENLVADLFYVFSLPILIFYFN 97
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 78 RRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVPHHVAVTISTRVAGIVC 137
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 98 KTDWIFGELCKLQRFIFVNLVYLSILFTCTISVHRYGVVHPMKSLGRLKK-SATIVC 156
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 138 T-LWALVILGTIVLLENHLCVQETAVS-CESFIMESANGWHDIMFQLEFF 186
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 157 VCVNITVMAGISPILYFSRTGLRRNKTNTCYDTTSKELLETFYFYSMTTF-----FGFC 211
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 187 MPLGIILFCSPKIVWSLRRLRQOLARQARKKATRIMVVAIVFTCYLP-----SVSA 239
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 212 IPFATILVCYGFIVKALISNDM--KTLRGRKSVRLVITLAVFAISYLPFHVKNLQ 269
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 240 RLYFLWTVPSSACDPSVHGALHITLFTYMNMLDPLVYFSSPSFPKYNKL--KICSL 297
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 270 RLYV-OGLDTCENWRRVYATQVTRGLASLNSCDVDPILYFLAGDTFRFRFTNAASRFMT 328
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 298 KPGQPGHKTORPEMPISNLGRSCISV 326
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 329 RSELSMQFRSEDSPLQPVSNISQNGDTS 357
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
Q90X57 PRELIMINARY; PRT; 361 AA.
ID Q90X57
AC Q90X57;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE P2Y1 NUCLEOTIDE RECEPTOR.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN
RP SEQUENCE FROM N.A.
RA Cheng A.W., Tsim K.W.;
RT "Cloning of Xenopus P2Y1 Receptor."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF432354; AAL27614.1; --
KW RECEPTOR.
SQ SEQUENCE 361 AA; 41002 MW; E5B2D605F5B57FED CRC64;

Query Match 18.4%; Score 341; DB 13; Length 361;
Best Local Similarity 29.0%; Pred. No. 6.4e-26;
Matches 93; Conservative 61; Mismatches 135; Indels 32; Gaps 9;

Qy 18 MPPLIIVAFVGLGNGVLCGFCFHMKTWKSTVYLFNLAADFLMLICLPFRDYLR 77
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 42 LPAVIVVTCITGFGNSVAINWNIHFMRPWSISIVMENLADFLYVLSLPALIFYFN 101
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 78 RRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVPHHVAVTISTRVAGIVC 137
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 102 KTDWIFGALCKLQRFIFVNLVYLSILFTCTISVHRYGVVHPMKSLGRLKKSIYISA 161
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 138 TLWALVILGTIVLLENHLCVQET-AVSC-----ESFIMESANGWHDIMFQLEFF 187
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 162 LWEIFVIAGISPILYFSRTGLRRNKTNTCTDTSSEYLSRVFYSM-----CTTVFGFCI 216
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```


THIS PAGE BLANK (USF...)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2002, 00:07:23 ; Search time 209 Seconds
(without alignments)
8551.713 Million cell updates/sec

Title: US-09-886-041-1
Perfect score: 1041
Sequence: 1 atgtacaacggctgtgctg.....acattgttgaggcactga 1041

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*				
1:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*			
2:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*			
3:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*			
4:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*			
5:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*			
6:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*			
7:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*			
8:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*			
9:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*			
10:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*			
11:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*			
12:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*			
13:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*			
14:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*			
15:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*			
16:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*			
17:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*			
18:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*			
19:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*			
20:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*			
21:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*			
22:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*			
23:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*			
24:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*			

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length	Description
1	1041	100.0	1041	22 AAS12581
2	1041	100.0	1041	22 AAS07946
3	1041	100.0	1041	22 AAS1008
4	1041	100.0	1050	22 ABA81529
5	1041	100.0	1050	22 ABA81530
6	1041	100.0	1083	24 AAD26371
7	1041	100.0	1730	24 AAS18501
8	1041	100.0	2331	22 AAS12582
9	1037.8	99.7	1104	22 ABA81531

cdNA encoding nove
Human cDNA encodin
Human nGPCR11 codi
Human GPCR1a polyn
Human GPCR1b polyn
Human G-protein co
cdNA encoding HM74
Gene encoding nove
Human GPCR1c polyn

10	888	85.3	888	22	AAS0974	Human nGPCR11 codi
11	372.4	35.8	1361	20	AAX16671	G-protein coupled
12	369.2	35.5	1092	22	AAS07952	Human cDNA encodin
13	367.6	35.3	1164	21	AAA30658	Human G protein-co
14	366	35.2	1164	21	AAA30738	DNA encoding human
15	220.2	21.2	546	15	AAQ55054	Sequence of orphan
16	188.8	18.1	1218	24	AAS98125	Human DNA for pote
17	188.8	18.1	1272	22	AAS07938	Human cDNA encodin
18	188.8	18.1	1317	22	AAI71609	Human oploid-type
19	188.8	18.1	1423	22	AAI82668	Human 7TM clone HE
20	188.8	18.1	1594	20	AAX06947	cdNA clone HE0AD54
21	188.8	18.1	1594	24	AAS98066	Human DNA for pote
22	188.8	18.1	1795	21	AAA27049	Human cell surface
23	188.8	18.1	1934	22	AAK84005	Human immune/hagma
24	188.8	18.1	2429	22	AAQ06502	Human CON103 G pro
25	188.8	18.1	4980	22	AAI71608	Human oploid-type
26	188	18.1	1435	20	AAX05948	HE0AD54 cDNA part1
27	184.8	17.8	8042	24	ABL32927	Human immune syste
28	181.2	17.4	8042	24	ABL32926	Human immune syste
29	140	13.4	9057	24	AAD26722	Human G-protein co
30	139.2	13.4	960	21	AAZ58587	Human G-protein co
31	139.2	13.4	960	24	AAD26666	Human G-protein co
32	139.2	13.4	9057	24	AAD26665	Human G-protein co
33	137.6	13.2	960	21	AAA30725	DNA encoding human
34	136	13.1	960	21	AAA30622	Human G protein-co
35	136	13.1	1488	22	ABA02183	Human G-protein co
36	109	10.5	1020	21	AAA30601	Human G protein-co
37	109	10.5	1901	15	AAQ66178	Seven transmembran
38	109	10.5	1901	19	AAV18356	Human R12 seven tr
39	109	10.5	1901	21	AAV91725	Human 7TM receptor
40	109	10.5	2453	18	AAT44092	Human G-protein th
41	105.8	10.2	1020	21	AAA30718	DNA encoding human
42	105.8	10.2	1098	21	AAA30740	DNA encoding human
43	104.2	10.0	1098	21	AAA30666	Human G protein-co
44	104.2	10.0	1281	23	AAS72718	DNA encoding novel
45	104.2	10.0	1281	23	AAS77014	DNA encoding novel

ALIGNMENTS

RESULT 1
AAS12581
ID AAS12581 standard; cdNA; 1041 BP.
XX
AC AAS12581;
XX
DT 19-DEC-2001 (first entry)
XX
DE cdNA encoding novel human G protein-coupled receptor (GPCR).
XX
DE Human; G-protein coupled receptor; GPCR; chemokine receptor; protease;
KW hyperproliferative disorder; neurological disorder; psychiatric disease;
KW inflammatory disorder; respiratory disorder; gene therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 1..1041
FT FT /*tag= a
FT FT /product= "GPCR protein"
XX
PN WO200173029-A2.
XX
PD 04-OCT-2001.
XX
PF 27-MAR-2001; 2001WO-US09522.
XX
PR 27-MAR-2000; 2000US-192419P.
PR 06-SEP-2000; 2000US-230459P.
PR 20-SEP-2000; 2000US-0666535.
XX
PA (PEKE) PE CORP NY.

XX
PI Ye J, Cravchik A, Di Francesco V, Beasley EM;
XX
WPI; 2001-61503/71.
DR P-PSDB; AAU06197.
DR

Novel human G-protein coupled receptor proteins and nucleic acid molecules encoding the protein for use in developing human therapeutics and diagnostic compositions and for identifying modulators of the protein

PS Claim 23; Fig 1; 66pp; English.

The present invention relates to the isolation of a novel human G-protein coupled receptor (GPCR) which is related to the chemokine receptor subfamily. The cDNA and gene sequences encoding for GPCR are also given in the invention. The sequences of the invention are useful for diagnosing and treating diseases or conditions mediated by human proteases. Such diseases include hyperproliferative disorders (e.g. hyperplasia), neurological disorders (e.g. Parkinson's disease), psychiatric diseases (e.g. schizophrenia), inflammatory disorders (e.g. diabetes) and respiratory disorders (e.g. adult respiratory distress syndrome, ARDS). The GPCR protein is also useful for identifying a modulator of the expression of the protein. It also serves as a target for identifying agents for use in mammalian therapeutic applications, e.g. a human drug, particularly modulating a biological or pathological response in a cell or tissue that expresses the protein, in biological assays related to GPCRs that are related to members of the chemokine receptor subfamily, in drug screening assays and in competition binding assays. GPCR is also useful in diagnosing a disease or predisposition to a disease mediated by the peptide, in pharmacogenomic analysis. The polynucleotide sequences can also be used in gene therapy. The present sequence encodes for the novel human GPCR of the invention.

Sequence 1041 BP; 208 A; 294 C; 278 G; 261 T; 0 other;

Query Match	100.0%	Score 1041;	DB 22;	Length 1041;
Best Local Similarity	100.0%;	Pred. No. 4e-288;		
Matches 1041;	Conservative	0;	Mismatches 0;	Indels 0;
Gaps	0;			

Qy	1	ATGTACAACGGGTGTCGTGCGCATCGAGGGGACACCATCTCCAGAGTGATGCGCGG	60
Db	1	ATGTACAACGGGTGTCGTGCGCATCGAGGGGACACCATCTCCAGAGTGATGCGCGG	60
Qy	61	CTGCTCATTTGTGSCCTTTGTGCTGGGCGCACATAGGCCAATGGGGTGGCCCTGTGCTGTTTC	120
Db	61	CTGCTCATTTGTGSCCTTTGTGCTGGGCGCACATAGGCCAATGGGGTGGCCCTGTGCTGTTTC	120
Qy	121	TGCTTCCACATGAAGACCTTGGGAAGCCAGCACTGTTTACCTTTTCAATTTGGCCCTGTGCT	180
Db	121	TGCTTCCACATGAAGACCTTGGGAAGCCAGCACTGTTTACCTTTTCAATTTGGCCCTGTGCT	180
Qy	181	GAATTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC	240
Db	181	GAATTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC	240
Qy	241	TGGGCTTTTGGGACACTTCCTCGCCAGATGGGGCTCTTACGCTTGCCCATGAACAGGGCC	300
Db	241	TGGGCTTTTGGGACACTTCCTCGCCAGATGGGGCTCTTACGCTTGCCCATGAACAGGGCC	300
Qy	301	GGGAGCATCGTGTTCCCTTACGGTGGTGGCTGGGACAGGTATTTCAAAGTGGTCCACCCC	360
Db	301	GGGAGCATCGTGTTCCCTTACGGTGGTGGCTGGGACAGGTATTTCAAAGTGGTCCACCCC	360
Qy	361	CACACGCGGTGAACACTATCTCCACCCGGGTGGCGGCTTGGCATCGTCTGCACCCCTGTGG	420
Db	361	CACACGCGGTGAACACTATCTCCACCCGGGTGGCGGCTTGGCATCGTCTGCACCCCTGTGG	420
Qy	421	GCCTTGGTCATCTCGGACAGTGTATCTTTTGTCTGGAGAACCATCTCTGCGCTGCAAGAG	480
Db	421	GCCTTGGTCATCTCGGACAGTGTATCTTTTGTCTGGAGAACCATCTCTGCGCTGCAAGAG	480
Qy	481	ACGGCGCTCTCCTGTGAGAGCTTCATCATGTGAGTCGGGCCAATGGGCTGGCATCACAATCATG	540

[illegible]

PR 16-NOV-1999; 99US-0165838.
PR 17-NOV-1999; 99US-0166071.
PR 19-NOV-1999; 99US-0166678.
PR 28-DEC-1999; 99US-0173396.
PR 22-FEB-2000; 2000US-0184129.
PR 28-FEB-2000; 2000US-0185421.
PR 28-FEB-2000; 2000US-0185554.
PR 02-MAR-2000; 2000US-0186530.
PR 03-MAR-2000; 2000US-0186811.
PR 09-MAR-2000; 2000US-0188114.
PR 17-MAR-2000; 2000US-0190310.
PR 21-MAR-2000; 2000US-0190800.
PR 20-APR-2000; 2000US-0198568.
PR 02-MAY-2000; 2000US-0201190.
PR 08-MAY-2000; 2000US-0203111.
PR 25-MAY-2000; 2000US-0207094.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
XX Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;
PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;
XX
XX WPI; 2001-389826/41.
DR P-PSDB; AAG80968.
XX
PT New G protein-coupled receptor (ngPCR-x) and its encoding
PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -
XX
XX Claim 4; Page 89; 261pp; English.
XX
CC The present invention relates to novel G protein-coupled receptors
CC (ngPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
CC 28, 31-38, 40, 41, 53-60) and their coding sequences. The present
CC sequence is the coding sequence for one such G protein-coupled receptor.
CC GPCRx are also known as seven transmembrane receptors and function in
CC signal transduction. The ngPCRx coding sequences are useful for
CC screening a human to diagnose a disorder affecting the brain or a genetic
CC predisposition, specifically schizophrenia. ngPCRx are useful for
CC identifying compounds useful for treating schizophrenia. Detection of
CC ngPCRx in a sample is useful as a diagnostic tool for diseases or
CC disorders e.g. thyroid disorders, renal failure, rheumatoid arthritis,
CC CNS disorders, infections such as HIV-1, metabolic and cardiovascular
CC diseases, proliferative disorders and hormonal disorders. Modulators of
CC ngPCRx activity have the utility for treating neurological disorders,
CC including schizophrenia, ADHD/ADD (attention deficit-hyperactivity
CC disorder/attention deficit disorder), and neuronal disorders such as
CC Alzheimer's disease, Parkinson's disease, migraine and senile dementia.
CC Additional disorders include inflammatory conditions (e.g. Crohn's
CC disease), rheumatoid arthritis, autoimmune disorders, cancers,
CC respiratory ailments such as asthma, and inflammatory diseases e.g.
CC inflammatory bowel disease.
XX
XX Sequence 1041 BP; 208 A; 294 C; 278 G; 261 T; 0 other;
XX
Query Match 100.0%; Score 1041; DB 22; Length 1041;
Best Local Similarity 100.0%; Pred. No. 4e-288;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGTACACAGGGTGTGTCGCGCATGAGGGGACACCATCTCCAGGTGATGCCGCG 60
Db 1 ATGTACACAGGGTGTGTCGCGCATGAGGGGACACCATCTCCAGGTGATGCCGCG 60
Qy 61 CTGCTCATTTGGCTTTGTCTGCTGGGCGCACTAGGCAATGGGTGCGCCCTGTGGTTTC 120
Db 61 CTGCTCATTTGGCTTTGTCTGCTGGGCGCACTAGGCAATGGGTGCGCCCTGTGGTTTC 120
Qy 121 TGCTTCCACATGAAGACCTGGAAGCCAGCAGCTGTTTACCTTTTCAATTTGGCGGTGGCT 180
Db 121 TGCTTCCACATGAAGACCTGGAAGCCAGCAGCTGTTTACCTTTTCAATTTGGCGGTGGCT 180
Qy 181 GATTTCCTCCTTATGATCTGCTGCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
Db 181 GATTTCCTCCTTATGATCTGCTGCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240

Qy 241 TGGGCTTTTGGGACATTCCTGCCAGTGGGGCTTTCACGTTGGCCATGAACAGGGCC 300
Db 241 TGGGCTTTTGGGACATTCCTGCCAGTGGGGCTTTCACGTTGGCCATGAACAGGGCC 300
Qy 301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTCGGACAGGTATTTCAAAGTGGTCCACCCC 360
Db 301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTCGGACAGGTATTTCAAAGTGGTCCACCCC 360
Qy 361 CACCACGGGTGAACACTATCTCCACCGGGTGGGGCTGGCATCGTCTGCACCTGTGG 420
Db 361 CACCACGGGTGAACACTATCTCCACCGGGTGGGGCTGGCATCGTCTGCACCTGTGG 420
Qy 421 GCCCTGGTCATCCTCGGAACAGTGTATCTTTTGTGGAGAACCATCTCTCGGTGCAAGAG 480
Db 421 GCCCTGGTCATCCTCGGAACAGTGTATCTTTTGTGGAGAACCATCTCTCGGTGCAAGAG 480
Qy 481 ACGGCGGTCTCTGTGAGAGCTTCATCATGAGTCGGCCAAATGGCTGGCATGACATCATG 540
Db 481 ACGGCGGTCTCTGTGAGAGCTTCATCATGAGTCGGCCAAATGGCTGGCATGACATCATG 540
Qy 541 TTCCAGCTGGAGTTCCTTTATGCCCTCGGCATCATCTTATTTTCTCTTCAAGATTGTT 600
Db 541 TTCCAGCTGGAGTTCCTTTATGCCCTCGGCATCATCTTATTTTCTCTTCAAGATTGTT 600
Qy 601 TGGAGCCTTGAGCGGAGGACAGCTGGCCAGACAGGCTCGGATGAAGAAGCGCACCCGG 660
Db 601 TGGAGCCTTGAGCGGAGGACAGCTGGCCAGACAGGCTCGGATGAAGAAGCGCACCCGG 660
Qy 661 TTTCATCATGTTGGTGGCAATTTGTTTCATCATGCTACCTGCTGCTGCTAGAG 720
Db 661 TTTCATCATGTTGGTGGCAATTTGTTTCATCATGCTACCTGCTGCTGCTAGAG 720
Qy 721 CTCATATTCCTCTCGAGCGTCCCTCGAGTGCCTCGCATCCCTCTGTCCATGGGCGCTG 780
Db 721 CTCATATTCCTCTCGAGCGTCCCTCGAGTGCCTCGCATCCCTCTGTCCATGGGCGCTG 780
Qy 781 CACATAACCCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTCGTGTATTTT 840
Db 781 CACATAACCCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTCGTGTATTTT 840
Qy 841 TCAAGCCCTCTCTTCCCAATTTCTACAAAGCTCAAAATCTGCAGTCTGAACCCCAAG 900
Db 841 TCAAGCCCTCTCTTCCCAATTTCTACAAAGCTCAAAATCTGCAGTCTGAACCCCAAG 900
Qy 901 CAGCCAGACACTCAAAACACAAAGCCGGAAGAGATGCCAATTTCCAACTCGGTGCG 960
Db 901 CAGCCAGACACTCAAAACACAAAGCCGGAAGAGATGCCAATTTCCAACTCGGTGCG 960
Qy 961 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGCAATGGATCCC 1020
Db 961 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGCAATGGATCCC 1020
Qy 1021 CACATTGTTAGTGGCACTGA 1041
Db 1021 CACATTGTTAGTGGCACTGA 1041

RESULT 4
ABAB1529
ID ABA81529 standard; DNA; 1050 BP.
XX
XX ABA81529;
XX
XX 28-JAN-2002 (first entry)
XX
XX Human GPCR1a polynucleotide SEQ ID NO 1.
DE Human; GPCR; G-coupled protein-receptor; cardiant; antiarteriosclerotic;
KW anabolic; cyostatic; antiviral; gene therapy; cardiomyopathy; obesity;
KW anorexia; diabetes; osteoporosis; Crohn's disease; multiple sclerosis;
KW asthma; Alzheimer's disease; Parkinson's disease; Huntington's disease;
KW infection; human immunodeficiency virus; HIV; ds.

XX Homo sapiens.
OS WO200174904-A2.
PN 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US10241.
XX 31-MAR-2000; 2000US-193664P.
PR 05-APR-2000; 2000US-194614P.
PR 06-APR-2000; 2000US-195063P.
PR 06-APR-2000; 2000US-195066P.
PR 06-APR-2000; 2000US-195067P.
PR 06-APR-2000; 2000US-195068P.
PR 06-APR-2000; 2000US-195069P.
PR 06-APR-2000; 2000US-195070P.
PR 06-APR-2000; 2000US-195510P.
PR 21-JUL-2000; 2000US-219855P.
PR 27-JUL-2000; 2000US-221284P.
PR 28-JUL-2000; 2000US-221325P.
PR 11-AUG-2000; 2000US-224588P.
PR 11-OCT-2000; 2000US-239613P.
PR 18-JAN-2001; 2001US-262508P.
PR 23-JAN-2001; 2001US-263433P.
PR 23-JAN-2001; 2001US-263604P.
PR 30-JAN-2001; 2001US-265161P.
PR 29-MAR-2001; 2001US-0823172.
XX (CURA-) CURAGEN CORP.
PA
XX
XX Majumder K, Vernet CAM, Casman SJ, Wolenc AR, Spaderna SK;
PI Padigaru M, Mishnu VS, Tchernev VT, Spytek KA, Li L;
PI Baumgartner JC, Gusev V;
XX
XX WPI: 2001-639351/73.
DR P-PSDB; ABB44522.
XX
XX New human G-protein coupled receptor X, GPCR_X, polypeptide useful in
PT treatment or prevention of GPCR_X associated disorders e.g.
PT cardiomyopathy or atherosclerosis, and to screen for antagonists and
PT agonists useful therapeutically -
XX
XX Claim 9; Page 7; 157pp; English.
XX
XX The invention relates to nucleic acid sequences (ABA81529-ABA81552) that
CC encode G-coupled protein-receptor related polypeptides
CC (ABB44522-ABB44543). The isolated polypeptide having a sequence differing
CC by no more than 15 % of amino acid residues from one of 22 amino acid
CC sequences (or mature forms of the sequences), fully defined in the
CC specification and corresponding to human G-protein coupled receptor X
CC (GPCR_X) polypeptides. The polypeptides have potential cardiac
CC antiarrhythmogenic, anabolic, cytostatic and antiviral activity. The
CC polypeptides can be administered therapeutically, especially using gene
CC therapy and expressing the encoding DNA in vivo, to treat or prevent
CC GPCR_X-associated disorders, especially in humans. For example, they can
CC be used to treat/prevent cardiomyopathy, atherosclerosis, disorders
CC related to signal processing and metabolic pathway modulation (e.g.
CC obesity, anorexia, diabetes, osteoporosis, Crohn's disease, multiple
CC sclerosis, asthma, cancers, neurodegenerative disorders (e.g. Alzheimer's
CC disease, Parkinson's disorder, Huntington's disease), immune disorders,
CC haematopoietic disorders, developmental diseases, neurological disorders,
CC bacterial, fungal, protozoal and viral infections (e.g. with human
CC immunodeficiency virus (HIV)-1 or HIV-2). They can be used diagnostically
CC to determine the presence of or predisposition to a disease associated
CC with altered levels of the polypeptide in mammals (especially humans) by
CC detecting alterations in polypeptide expression levels relative to
CC control samples. They are useful to identify agents binding polypeptide
CC (e.g. cellular receptors or downstream effectors) and/or agents
CC modulating cellular polypeptide expression or activity, useful as
CC antagonists and agonists in disease treatment.
XX
XX Sequence 1050 BP; 211 A; 299 C; 279 G; 261 T; 0 other;

Query Match 100.0%; Score 1041; DB 22; Length 1050;
Best Local Similarity 100.0%; Pred. No. 4e-288;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTACAACGGGTGCTGCTGCCGATCGAGGGGAGACACCATCTCCAGGTGATGCGCGCG 60
DB 6 ATGTACAACGGGTGCTGCTGCCGATCGAGGGGAGACACCATCTCCAGGTGATGCGCGCG 65
QY 61 CTGCTCATTTGTGGCCTTTTGCTGGGGCAGTGGCAATGGGCTGCCCTGTGTGGTTTC 120
DB 66 CTGCTCATTTGTGGCCTTTTGCTGGGGCAGTGGCAATGGGCTGCCCTGTGTGGTTTC 125
QY 121 TGTCTTCACATGAAGACCTGGAAGCCGACGACTGTTTACATTTTCAATTTGGCCGTGGCT 180
DB 126 TGTCTTCACATGAAGACCTGGAAGCCGACGACTGTTTACATTTTCAATTTGGCCGTGGCT 185
QY 181 GATTTCTCTTATGATGCTGCTTTCGCTGCTTTCGACAGACTATTACCTCAGACGTAGACAC 240
DB 186 GATTTCTCTTATGATGCTGCTTTCGCTGCTTTCGACAGACTATTACCTCAGACGTAGACAC 245
QY 241 TGGGCTTTTGGGACATTCCTCGCGAGTGGGCTCTTACGTTGGCCATGAACAGGGCC 300
DB 246 TGGGCTTTTGGGACATTCCTCGCGAGTGGGCTCTTACGTTGGCCATGAACAGGGCC 305
QY 301 GGGAGCATCGTGTCTTACGCTGGTGGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC 360
DB 306 GGGAGCATCGTGTCTTACGCTGGTGGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC 365
QY 361 CACCACGCGGTGAACACTATCTCCACCCGGTGGCGCTGGCATGCTGCACCCCTGTGG 420
DB 366 CACCACGCGGTGAACACTATCTCCACCCGGTGGCGCTGGCATGCTGCACCCCTGTGG 425
QY 421 GCCTTGGTCATCTCTGGGAACAGTGTATCTTTTGTCTGGAGAACCATCTCTCGGTGCAAGAG 480
DB 426 GCCTTGGTCATCTCTGGGAACAGTGTATCTTTTGTCTGGAGAACCATCTCTCGGTGCAAGAG 485
QY 481 ACGGCGCTCTCTGTGAGAGCTTTCATCATGGAGTGGGCCAATGGCTGGCATGACATCATG 540
DB 486 ACGGCGCTCTCTGTGAGAGCTTTCATCATGGAGTGGGCCAATGGCTGGCATGACATCATG 545
QY 541 TTCAGCTGGAGTCTTTTATGCCCCCTCGGCATCATCTTATTTTGTCTTCAAGATTGTT 600
DB 546 TTCAGCTGGAGTCTTTTATGCCCCCTCGGCATCATCTTATTTTGTCTTCAAGATTGTT 605
QY 601 TGGAGCTGAGCGGAGGAGAGCTGGCCAGACAGCTCGGATGAGAGAGGAGCCCGG 660
DB 606 TGGAGCTGAGCGGAGGAGAGCTGGCCAGACAGCTCGGATGAGAGAGGAGCCCGG 665
QY 661 TTCATCATGTTGTTGGCAATTGTGTTTCATCACATGCTTACCTGCCCGAGCTGTCTGTAGA 720
DB 666 TTCATCATGTTGTTGGCAATTGTGTTTCATCACATGCTTACCTGCCCGAGCTGTCTGTAGA 725
QY 721 CTCTATTTCTTGGAGGGTGGCCCTCGAGTGCCTGGATCCCTCTGTCATGAGGGCCCTG 780
DB 726 CTCTATTTCTTGGAGGGTGGCCCTCGAGTGCCTGGATCCCTCTGTCATGAGGGCCCTG 785
QY 781 CACATACCTCTAGCTTTCACCTACATGAACAGACTGCTGATCCCTGGTGTATTATTTT 840
DB 786 CACATACCTCTAGCTTTCACCTACATGAACAGACTGCTGATCCCTGGTGTATTATTTT 845
QY 841 TCAAGCCCTCTCTTTTCCCAAAATCTACAACAGCTCAAAATCTGCAGTCTGAAACCCCAAG 900
DB 846 TCAAGCCCTCTCTTTTCCCAAAATCTACAACAGCTCAAAATCTGCAGTCTGAAACCCCAAG 905
QY 901 CAGCCAGGACACTCAAAAACACAAAGCCCGGAGAGATGCAATTTTCGAACCTCGGTGCG 960
DB 906 CAGCCAGGACACTCAAAAACACAAAGCCCGGAGAGATGCAATTTTCGAACCTCGGTGCG 965
QY 961 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGCAATGGATCCC 1020
DB 966 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGCAATGGATCCC 1025

Qy 1021 CACATTGTTAGTGGCACTGA 1041
 Db 1026 CACATTGTTAGTGGCACTGA 1046

RESULT 5
 ID ABA81530 standard; DNA: 1050 BP.
 AC ABA81530;
 XX ABA81530;

DT 28-JAN-2002 (first entry)
 DE Human GPCR1b polynucleotide SEQ ID NO 3.
 XX Human: GPCR; G-coupled protein-receptor; cardiant; antiarteriosclerotic;

KW anabolic; cytosolic; antiviral; gene therapy; cardiomyopathy; obesity;
 KW anorexia; diabetes; osteoporosis; Crohn's disease; multiple sclerosis;
 KW asthma; Alzheimer's disease; Parkinson's disorder; Huntington's disease;
 KW infection; human immunodeficiency virus; HIV; ds.
 XX Homo sapiens.

OS WO200174904-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US10241.

XX 31-MAR-2000; 2000US-193664P.

XX 05-APR-2000; 2000US-194614P.

XX 06-APR-2000; 2000US-195063P.

XX 06-APR-2000; 2000US-195066P.

XX 06-APR-2000; 2000US-195067P.

XX 06-APR-2000; 2000US-195068P.

XX 06-APR-2000; 2000US-195069P.

XX 06-APR-2000; 2000US-195510P.

XX 21-JUL-2000; 2000US-219855P.

XX 27-JUL-2000; 2000US-221284P.

XX 28-JUL-2000; 2000US-221325P.

XX 11-AUG-2000; 2000US-224588P.

XX 11-OCT-2000; 2000US-239613P.

XX 18-JAN-2001; 2001US-262508P.

XX 23-JAN-2001; 2001US-263433P.

XX 30-JAN-2001; 2001US-263604P.

XX 29-MAR-2001; 2001US-265161P.

XX (CURA-) CURAGEN CORP.

XX Majumder K, Vernet CAM, Casman SJ, Wolenc AR, Spaderna SK;

XX Padigar M, Mishnu VS, Tchernev VT, Spytek KA, Li L;

XX Baumgarthner JC, Gusev VY;

XX WPI: 2001-639351/73.

XX P-PSDB: ABB44522.

XX New human G-protein coupled receptor X, GPCR, polypeptide useful in

XX treatment or prevention of GPCR associated disorders e.g.

XX cardiomyopathy or arteriosclerosis, and to screen for antagonists and

XX agonists useful therapeutically -

XX Claim 9; Page 10; 157pp; English.

CC polypeptides can be administered therapeutically, especially using gene
 CC therapy and expressing the encoding DNA in vivo, to treat or prevent
 CC GPCR-associated disorders, especially in humans. For example, they can
 CC be used to treat/prevent cardiomyopathy, arteriosclerosis, disorders
 CC related to signal processing and metabolic pathway modulation (e.g.
 CC obesity, anorexia), diabetes, osteoporosis, Crohn's disease, multiple
 CC sclerosis, asthma, cancers, neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Parkinson's disorder, Huntington's disease), immune disorders,
 CC haematopoietic disorders, developmental diseases, neurological disorders,
 CC bacterial, fungal, protozoal and viral infections (e.g. with human
 CC immunodeficiency virus (HIV)-1 or HIV-2). They can be used diagnostically
 CC to determine the presence of or predisposition to a disease associated
 CC with altered levels of the polypeptide in mammals (especially humans) by
 CC detecting alterations in polypeptide expression levels relative to
 CC control samples. They are useful to identify agents binding polypeptide
 CC (e.g. cellular receptors or downstream effectors) and/or agents
 CC modulating cellular polypeptide expression or activity, useful as
 CC antagonists and agonists in disease treatment.
 XX
 SQ Sequence 1050 BP; 211 A; 298 C; 279 G; 262 T; 0 other;

Query Match 100.0%; Score 1041; DB 22; Length 1050;

Best Local Similarity 100.0%; Pred. No. 4e-288;

Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTACAAACGGTCTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 60

Db 6 ATGTACAAACGGTCTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 65

Qy 61 CTGCTCATTTGGGCTTTGCTGGGCGCACTAGGCAATGGGGTGGCCCTGTGGTTTC 120

Db 66 CTGCTCATTTGGGCTTTGCTGGGCGCACTAGGCAATGGGGTGGCCCTGTGGTTTC 125

Qy 121 TGCTTCCACATGAAGACCTGGAAGCCGAGCACTGTTTACCTTTTCAATTTGGCGTGGCT 180

Db 126 TGCTTCCACATGAAGACCTGGAAGCCGAGCACTGTTTACCTTTTCAATTTGGCGTGGCT 185

Qy 181 GATTTCTCTTATGATCTGCTGCTTTCGGACAGACTATTACCTCAGACGTAGACAC 240

Db 186 GATTTCTCTTATGATCTGCTGCTTTCGGACAGACTATTACCTCAGACGTAGACAC 245

Qy 241 TGGGCTTTTGGGACATTCCTGCGAGTGGGCTTTCAGTTGGCCATGAACAGGGCC 300

Db 246 TGGGCTTTTGGGACATTCCTGCGAGTGGGCTTTCAGTTGGCCATGAACAGGGCC 305

Qy 301 GGGAGCATCGTGTTCCTTACCGTGTGCTCGGACAGGTATTTCAAAGTGGTCCACCCC 360

Db 306 GGGAGCATCGTGTTCCTTACCGTGTGCTCGGACAGGTATTTCAAAGTGGTCCACCCC 365

Qy 361 CACCACGGGTGAACACTATCTCCACCGGGTGGCGGTGGCATCGTGTGCACCTGTGG 420

Db 366 CACCACGGGTGAACACTATCTCCACCGGGTGGCGGTGGCATCGTGTGCACCTGTGG 425

Qy 421 GCCCTGGTCACTCTGGGAACAGTGTATCTTTTGGTGGAGAACCATCTCTCGGTGCAAGAG 480

Db 426 GCCCTGGTCACTCTGGGAACAGTGTATCTTTTGGTGGAGAACCATCTCTCGGTGCAAGAG 485

Qy 481 ACGCCGCTCTCTCTGAGAGCTTCATCATGAGTCGGCAATGGTGGCATGACATCATG 540

Db 486 ACGCCGCTCTCTCTGAGAGCTTCATCATGAGTCGGCAATGGTGGCATGACATCATG 545

Qy 541 TTCCAGCTGGAGTTCCTTTATGCCCTCGGCATCATCTTATTTTCTCTCTCAAGATTGTT 600

Db 546 TTCCAGCTGGAGTTCCTTTATGCCCTCGGCATCATCTTATTTTCTCTCTCAAGATTGTT 605

Qy 601 TGGAGCTTGAGCGGAGGACAGCAGCTGGCCAGACAGGCTCGGATGAAGAGGCCGCCGG 660

Db 606 TGGAGCTTGAGCGGAGGACAGCAGCTGGCCAGACAGGCTCGGATGAAGAGGCCGCCGG 665

Qy 661 TTCATCATGGTGGCAATTTGTTTCATCATCATGCTTACCTGCCAGCGTCTCTGCTAGA 720

Db 666 TTCATCATGGTGGCAATTTGTTTCATCATCATGCTTACCTGCCAGCGTCTCTGCTAGA 725

CC The invention relates to nucleic acid sequences (ABA81529-ABA81552) that
 CC encode G-coupled protein-receptor related polypeptides
 CC (ABB44522-ABB44543). The isolated polypeptide having a sequence differing
 CC by no more than 15 % of amino acid residues from one of 22 amino acid
 CC sequences (or mature forms of the sequences), fully defined in the
 CC specification and corresponding to human G-protein coupled receptor X
 CC (GPCR) polypeptides. The polypeptides have potential cardiant,
 CC antiarteriosclerotic, anabolic, cytostatic and antiviral activity. The


```
Qy 421 GCCCTGTCATCCTGGGAACAGTGTATCTTTTGTGGAGAACCATCTCTCGGTGCAAGAG 480
Db 463 GCCCTGTCATCCTGGGAACAGTGTATCTTTTGTGGAGAACCATCTCTCGGTGCAAGAG 522
Qy 481 ACGCCCTCTCTCTGTAGAGCTTTCATCATGAGAGTCGGCAATGCTGGCATGACATCATG 540
Db 523 ACGCCCTCTCTCTGTAGAGCTTTCATCATGAGAGTCGGCAATGCTGGCATGACATCATG 582
Qy 541 TTCCAGCTGGAGTCTTTATGCCCTCGGCATCATCTTATTTTCTCTCTCAAGATTGTT 600
Db 583 TTCCAGCTGGAGTCTTTATGCCCTCGGCATCATCTTATTTTCTCTCTCAAGATTGTT 642
Qy 601 TGGAGCCTGAGCGGAGCGACAGCTGGCCACAGAGCTCGGATGAAGAGCGGCCCGG 660
Db 643 TGGAGCCTGAGCGGAGCGACAGCTGGCCACAGAGCTCGGATGAAGAGCGGCCCGG 702
Qy 661 TTATCATGTGGTGGCAATTTGTTTTCATCATGCTACCTGCGCCAGGGTGTCTGTAGTA 720
Db 703 TTATCATGTGGTGGCAATTTGTTTTCATCATGCTACCTGCGCCAGCGTGTCTGTAGTA 762
Qy 721 CTCTATTTCCTCTGAGCGTCCCTCGAGTGCCTGCGATCCCTCTGTCCATGGGCGCTG 780
Db 763 CTCTATTTCCTCTGAGCGTCCCTCGAGTGCCTGCGATCCCTCTGTCCATGGGCGCTG 822
Qy 781 CACATAACCCCTCAGCTTACCTACATGAACAGCATGCTGGATCCCTCGGTGTTATTTT 840
Db 823 CACATAACCCCTCAGCTTACCTACATGAACAGCATGCTGGATCCCTCGGTGTTATTTT 882
Qy 841 TCAAGCCCTCTCTTCCCAATTTACACAAAGCTCAAAATCTCAGTCTGAACCCCAAG 900
Db 883 TCAAGCCCTCTCTTCCCAATTTACACAAAGCTCAAAATCTCAGTCTGAACCCCAAG 942
Qy 901 CAGCAGGACACTCAAAACAAAGGCGGAAGAGATGCCAATTCGAACCTCGGTGCG 960
Db 943 CAGCAGGACACTCAAAACAAAGGCGGAAGAGATGCCAATTCGAACCTCGGTGCG 1002
Qy 961 AGGAGTTCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGCAATGGATCCC 1020
Db 1003 AGGAGTTCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGCAATGGATCCC 1062
Qy 1021 CACATTGTTGAGTGGCACTGA 1041
Db 1063 CACATTGTTGAGTGGCACTGA 1083
```

RESULT 7

AA18501
ID AA18501 standard; cDNA; 1730 BP.

AC AA18501;

XX 26-FEB-2002 (first entry)

DT 26-FEB-2002 (first entry)

XX cDNA encoding HM74-like G-protein coupled receptor (GPCR).

DE HM74-like GPCR; G-protein coupled receptor; antibacterial; fungicide;
XX protozoacide; analgesic; cytostatic; neuroleptic; nootropic;
KW anticonvulsant; tranquilizer; viral infection; pain; cancer; anorexia;
KW bulimia; asthma; central nervous system disease; CNS disease;
KW cardiovascular disease; hypotension; hypertension; angina pectoris;
KW myocardial infarction; urinary retention; osteoporosis; ulcer; asthma;
KW inflammation; allergy; benign prostatic hypertrophy; multiple sclerosis;
KW psychotic disorder; neurological disorder; dyskinesia;
KW Huntington's disease; Tourette's syndrome; anxiety; schizophrenia;
KW manic depression; delirium; dementia; mental retardation; ss.

XX Homo sapiens.

XX Location/Qualifiers

FT 464..1504

FT CDS /*tag= a

FT /product= "HM74-like_GPCR"

PT /note= "G-protein coupled receptor"

XX WO200177320-A2.

XX 18-OCT-2001.

XX 04-APR-2001; 2001WO-EP03811.

XX 05-APR-2000; 2000US-194701P.

XX (FARB) BAYER AG.

XX Xiao Y;

XX WPI; 2002-049147/06.

XX P-PSDB; AAU11401.

XX Novel isolated polynucleotide, useful for treating infection, pain,
PT cancer, asthma, hypertension, myocardial infarction, urinary retention,
PT osteoporosis, encodes the human HM74-like G-protein coupled receptor
PT polypeptide -

XX Claim 1; Fig 1; 77pp; English.

XX The invention describes a novel isolated polynucleotide (I) encoding a
CC human HM74-like G-protein coupled receptor (GPCR) polypeptide. Reagents
CC that regulate HM74-like GPCR are useful for modulating the activity of
CC the protein in a disease selected from bacterial, fungal, protozoan, and
CC viral infection, pain, cancer, anorexia, bulimia, asthma, central nervous
CC system (CNS) disease, cardiovascular disease, hypotension, hypertension,
CC angina pectoris, myocardial infarction, urinary retention, osteoporosis,
CC ulcer, asthma, inflammation, allergy, benign prostatic hypertrophy,
CC multiple sclerosis and dyskinesia such as Huntington's disease and
CC Tourette's syndrome. The composition is also useful for treating
CC psychotic and neurological disorders such as anxiety, schizophrenia,
CC manic depression, delirium, dementia and severe mental retardation. (I)
CC or the HM74-like GPCR polypeptide are also useful for treating the above
CC mentioned diseases. (I) is useful in a diagnostic assay for detecting
CC diseases, susceptibility to diseases and abnormalities related to the
CC presence of mutations in the nucleic acid sequences which encode a GPCR.
CC The polypeptide is useful to identify test compounds which may act as
CC agonists or antagonists at the receptor site and which can be regulated
CC to provide therapeutic effects. The polypeptide is also useful as a bait
CC protein in a two-hybrid or three-hybrid assay, and to immunise a mammal
CC for production of polyclonal antibodies. This sequence encodes the human
CC HM74-like GPCR described in the method of the invention.

XX Sequence 1730 BP; 361 A; 494 C; 444 G; 431 T; 0 other;

XX Query Match 100.0%; Score 1041; DB 24; Length 1730;

XX Best Local Similarity 100.0%; Pred. No. 5.1e-288;

XX Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTACAAACGGTCTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATCCGCCCG 60

Db 464 ATGTACAAACGGTCTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATCCGCCCG 523

Qy 61 CTGCTCATTTGGGCTTTGTGCTGGCGCACTAGGCAATGGGTCGCCCTGTGTGTTTC 120

Db 524 CTGCTCATTTGGGCTTTGTGCTGGCGCACTAGGCAATGGGTCGCCCTGTGTGTTTC 583

Qy 121 TGCTTCCACATGAAGACCTGGAAGCCCGACGACTGTTTACCTTTTCAATTTGGCGTGGCT 180

Db 584 TGCTTCCACATGAAGACCTGGAAGCCCGACGACTGTTTACCTTTTCAATTTGGCGTGGCT 643

Qy 181 GATTTCCTCTTATGATCTGCTGCCCTTTCCGACACACTATTACCTCAGACGTAGACAC 240

Db 644 GATTTCCTCTTATGATCTGCTGCCCTTTCCGACACACTATTACCTCAGACGTAGACAC 703

Qy 241 TGGGCTTTTGGGGACATTCCCTGCGAGTGGGCTCTTACGTTGGGCATGAACAGGGCC 300

Db 704 TGGGCTTTTGGGGACATTCCCTGCGAGTGGGCTCTTACGTTGGGCATGAACAGGGCC 763

```
QY 301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGGGACAGAGTATTTCAAAGTGTCTCCACCCC 360
D 764 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGGGACAGAGTATTTCAAAGTGTCTCCACCCC 823
QY 361 CACCACGGGTGAACACTATCTCCACCGGGTGGCGGTGGCATCGTCTGCACCCCTGTGG 420
D 824 CACCACGGGTGAACACTATCTCCACCGGGTGGCGGTGGCATCGTCTGCACCCCTGTGG 883
QY 421 GCCTGTGTATCTCTGGGAACAGTGTATCTTTTGTCTGGAGAACCATCTCTGCGTGAAGAG 480
D 884 GCCTGTGTATCTCTGGGAACAGTGTATCTTTTGTCTGGAGAACCATCTCTGCGTGAAGAG 943
QY 481 ACGGCCGTCTCTGTGAGAGCTTCATCATGGAGTGGCCCAATGGCTGGCATGACATCATG 540
D 944 ACGGCCGTCTCTGTGAGAGCTTCATCATGGAGTGGCCCAATGGCTGGCATGACATCATG 1003
QY 541 TTCCAGCTGGAGTCTTTATGGCCCTCGGCATCATCTTATTTTGTCTCTTCAAGATTGTT 600
D 1004 TTCCAGCTGGAGTCTTTATGGCCCTCGGCATCATCTTATTTTGTCTCTTCAAGATTGTT 1063
QY 601 TGGAGCTTGAGCGGAGGACAGCTGGCCAGACAGCTCGGATGAAGAGGCGACCCGG 660
D 1064 TGGAGCTTGAGCGGAGGACAGCTGGCCAGACAGCTCGGATGAAGAGGCGACCCGG 1123
QY 661 TTCATCATGGTGTGGCAATTGTGTTTCATCATGCTACCTGCCCCAGCGTGTCTGCTAGA 720
D 1124 TTCATCATGGTGTGGCAATTGTGTTTCATCATGCTACCTGCCCCAGCGTGTCTGCTAGA 1183
QY 721 CTCATTTCTCTCTGGAGGTGCCCTGGAGTGGCTGGATCCCTCTGTCATGGGGCCCTG 780
D 1184 CTCATTTCTCTCTGGAGGTGCCCTGGAGTGGCTGGATCCCTCTGTCATGGGGCCCTG 1243
QY 781 CACATAACCTCTAGCTTACCTACATGAACAGCATGCTGGATCCCTCTGGTGTATTATTTT 840
D 1244 CACATAACCTCTAGCTTACCTACATGAACAGCATGCTGGATCCCTCTGGTGTATTATTT 1303
QY 841 TCAAGCCCTCTCTTCCCAATTTCTACAAAGCTCAAAATCTGCAGTCTGAACCCCAAG 900
D 1304 TCAAGCCCTCTCTTCCCAATTTCTACAAAGCTCAAAATCTGCAGTCTGAACCCCAAG 1363
QY 901 CAGCCAGGACACTCAAAACACAAAGCCGGAAGAGATGCCAATTTTCGAACCTCGGTCCG 960
D 1364 CAGCCAGGACACTCAAAACACAAAGCCGGAAGAGATGCCAATTTTCGAACCTCGGTCCG 1423
QY 961 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGCAATGGGATCCC 1020
D 1424 AGGAGTTGCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGCAATGGGATCCC 1483
QY 1021 CACATTGTTGAGTGGCACTGA 1041
D 1484 CACATTGTTGAGTGGCACTGA 1504

RESULT 8
AAS12582
ID AAS12582 standard; DNA; 2331 BP.
XX AC
AC AAS12582;
XX
XX
XX 19-DEC-2001 (first entry)
DE
DE Gene encoding novel human G protein-coupled receptor (GPCR).
XX
XX Human; G-protein coupled receptor; GPCR; chemokine receptor; protease;
KW hyperproliferative disorder; neurological disorder; psychiatric disease;
KW inflammatory disorder; respiratory disorder; gene therapy; ds.
XX
XX Homo sapiens.
XX
XX WO200173029-A2.
XX
XX 04-OCT-2001.
XX
```

```
PF 27-MAR-2001; 2001WO-US09522.
XX
PR 27-MAR-2000; 2000US-192419P.
PR 06-SEP-2000; 2000US-230459P.
XX 20-SEP-2000; 2000US-066653S.
PA (PERE ) PE CORP NY.
XX
PI Ye J, Cravchik A, Di Francesco V, Beasley EM;
DR WPI; 2001-616503/71.
XX
PT Novel human G-protein coupled receptor proteins and nucleic acid
PT molecules encoding the protein for use in developing human therapeutics
PT and diagnostic compositions and for identifying modulators of the
PT protein.
XX
PS Claim 23; Fig 3; 66pp; English.
XX
CC The present invention relates to the isolation of a novel human G-protein
CC coupled receptor (GPCR) which is related to the chemokine receptor
CC subfamily. The cDNA and gene sequences encoding for GPCR are also
CC given in the invention. The sequences of the invention are useful
CC for diagnosing and treating diseases or conditions mediated by human
CC proteases. Such diseases include hyperproliferative disorders
CC (e.g. hyperplasia), neurological disorders (e.g. Parkinson's disease),
CC psychiatric diseases (e.g. schizophrenia), inflammatory disorders
CC (e.g. diabetes) and respiratory disorders (e.g. adult respiratory
CC distress syndrome, ARDS). The GPCR protein is also useful for identifying
CC a modulator of the expression of the protein. It also serves as a target
CC for identifying agents for use in mammalian therapeutic applications,
CC e.g. a human drug, particularly modulating a biological or pathological
CC response in a cell or tissue that expresses the protein, in biological
CC assays related to GPCRs that are related to members of the chemokine
CC receptor subfamily, in drug screening assays and in competition binding
CC assays. GPCR is also useful in diagnosing a disease or predisposition to
CC a disease mediated by the peptide, in pharmacogenomic analysis. The
CC polynucleotide sequences can also be used in gene therapy. The present
CC sequence represents the human GPCR gene sequence of the invention.
XX
SQ Sequence 2331 BP; 497 A; 620 C; 592 G; 622 T; 0 other;

Query Match 100.0%; Score 1041; DB 22; Length 2331;
Best Local Similarity 100.0%; Pred. No. 5.9e-288;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTACAACGGGTGCTGCTGCCCATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 60
D 311 ATGTACAACGGGTGCTGCTGCCCATCGAGGGGACACCATCTCCAGGTGATGCCGCCG 370
QY 61 CTGCTCATTTGGCCCTTTGTGCTGGCGCACTAGGCAATGGGCTGCGCCCTGTGTGTTTC 120
D 371 CTGCTCATTTGGCCCTTTGTGCTGGCGCACTAGGCAATGGGCTGCGCCCTGTGTGTTTC 430
QY 121 TGCTTCCACATGAAGACCTGGAAGCCAGCACCTGTTTACCTTTTCAATTTGGCCGTGGCT 180
D 431 TGCTTCCACATGAAGACCTGGAAGCCAGCACCTGTTTACCTTTTCAATTTGGCCGTGGCT 490
QY 181 GATTTCCTCTTATGATCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
D 491 GATTTCCTCTTATGATCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 550
QY 241 TGGGCTTTTGGGACATTCCTCCGAGTGGGCTCTTACGTTGGCCATGAACAGGCC 300
D 551 TGGGCTTTTGGGACATTCCTCCGAGTGGGCTCTTACGTTGGCCATGAACAGGCC 610
QY 301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGAGTATTTCAAAGTGTCTCCACCCC 360
D 611 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGAGTATTTCAAAGTGTCTCCACCCC 670
QY 361 CACCACGGGTGAACACTATCTCCACCCGGTGGCGCTGGCATGCTGTCGACCCCTGTGG 420
D 671 CACCACGGGTGAACACTATCTCCACCCGGTGGCGCTGGCATGCTGTCGACCCCTGTGG 730
```


QY 421 GCCCTGTCATCCTGGGAACAGTGTATCTTTTGTGGAGAACCACTCTCTCGTGCACAGAG 480
DB 731 GCCCTGTCATCCTGGGAACAGTGTATCTTTTGTGGAGAACCACTCTCTCGTGCACAGAG 790
QY 481 ACGGCCGTCCTCTGTAGAGCTTTCATCATGAGAGTCGGCAATGCTGCATGACATCATG 540
DB 791 ACGGCCGTCCTCTGTAGAGCTTTCATCATGAGAGTCGGCAATGCTGCATGACATCATG 850
QY 541 TTCCAGCTGGAGTCTTTATGCCCTCGGCATCATCTATTATTTCCTCTCTCAAGATTGTT 600
DB 851 TTCCAGCTGGAGTCTTTATGCCCTCGGCATCATCTATTATTTCCTCTCTCAAGATTGTT 910
QY 601 TGGAGCTGGAGCGGAGCGACAGCTGCCACAGAGCTCGGATGAAGAGCGGACCCCGG 660
DB 911 TGGAGCTGGAGCGGAGCGACAGCTGCCACAGAGCTCGGATGAAGAGCGGACCCCGG 970
QY 661 TTCATCATGTTGGTGGCAATTTGTTTCATCATGCTACCTGCCAGCGGTCTGCTGAGA 720
DB 971 TTCATCATGTTGGTGGCAATTTGTTTCATCATGCTACCTGCCAGCGGTCTGCTGAGA 1030
QY 721 CTCATTTCTCTGAGCGTCCCTCGAGTCCCTCGATCCCTCTGTCATGGGCGCTG 780
DB 1031 CTCATTTCTCTGAGCGTCCCTCGAGTCCCTCGATCCCTCTGTCATGGGCGCTG 1090
QY 781 CACATAACCTTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTCGTGTATTATTTT 840
DB 1091 CACATAACCTTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTCGTGTATTATTTT 1150
QY 841 TCAAGCCCTCTCTTCCCAATTTACACCAAGCTCAAAATCTCAGTCTGAACCCCAAG 900
DB 1151 TCAAGCCCTCTCTTCCCAATTTACACCAAGCTCAAAATCTCAGTCTGAACCCCAAG 1210
QY 901 CAGCAGGACACTCAAAACAAAGCGGAGAGATGCCAATTTGCAACTCGGTGCG 960
DB 1211 CAGCAGGACACTCAAAACAAAGCGGAGAGATGCCAATTTGCAACTCGGTGCG 1270
QY 961 AGGAGTTCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGCAATGGATCCC 1020
DB 1271 AGGAGTTCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGATGGCAATGGATCCC 1330
QY 1021 CACATTTGTAGTGGCACTGA 1041
DB 1331 CACATTTGTAGTGGCACTGA 1351

RESULT 9

ABA81531

ID ABA81531 standard; DNA; 1104 BP.

AC ABA81531;

DT 28-JAN-2002 (first entry)

XX Human GPCR1c polynucleotide SEQ ID NO 4.

DE Human GPCR1c polynucleotide SEQ ID NO 4.

XX Human; GPCR; G-coupled protein-receptor; cardiant; antiarteriosclerotic;
KW anabolic; cytosolic; antiviral; gene therapy; cardiomyopathy; obesity;
KW anorexia; diabetes; osteoporosis; Crohn's disease; multiple sclerosis;
KW asthma; Alzheimer's disease; Parkinson's disorder; Huntington's disease;
KW infection; human immunodeficiency virus; HIV; ds.

XX Homo sapiens.

OS OS

XX WO200174904-A2.

PN PN

XX 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US10241.

XX 31-MAR-2000; 2000US-193664P.

PR 05-APR-2000; 2000US-194614P.

PR 06-APR-2000; 2000US-195063P.

PR 06-APR-2000; 2000US-195066P.
PR 06-APR-2000; 2000US-195067P.
PR 06-APR-2000; 2000US-195068P.
PR 06-APR-2000; 2000US-195069P.
PR 06-APR-2000; 2000US-195070P.
PR 06-APR-2000; 2000US-195510P.
PR 21-JUL-2000; 2000US-219855P.
PR 27-JUL-2000; 2000US-221284P.
PR 28-JUL-2000; 2000US-221325P.
PR 11-AUG-2000; 2000US-224588P.
PR 11-OCT-2000; 2000US-239613P.
PR 18-JAN-2001; 2001US-262598P.
PR 23-JAN-2001; 2001US-263433P.
PR 30-JAN-2001; 2001US-263604P.
PR 29-MAR-2001; 2001US-265161P.
XX 29-MAR-2001; 2001US-0823172.
PA (CURA-) CURAGEN CORP.
XX Majumder K, Vernet CAM, Casman SJ, Wolenc AR, Spaderna SK;
PI Padigaru M, Mishnu VS, Tchernev VT, Spytek KA, Li L;
PI Baumgartner JC, Gusev VV;
XX WPI; 2001-639351/73.
DR P-PSDB; ABB44523.
XX New human G-protein coupled receptor X, GPCR, polypeptide useful in
PT treatment or prevention of GPCR associated disorders e.g.
PT cardiomyopathy or atherosclerosis, and to screen for antagonists and
PT agonists useful therapeutically -
XX Claim 9; Page 11; 157pp; English.

XX The invention relates to nucleic acid sequences (ABA81529-ABA81552) that
CC encode G-coupled protein-receptor related polypeptides
CC (ABA44522-ABA44543). The isolated polypeptide having a sequence differing
CC by no more than 15 % of amino acid residues from one of 22 amino acid
CC sequences (or mature forms of the sequences), fully defined in the
CC specification and corresponding to human G-protein coupled receptor X
CC (GPCR) polypeptides. The polypeptides have potential cardiant,
CC antiarteriosclerotic, anabolic, cytostatic and antiviral activity. The
CC polypeptides can be administered therapeutically, especially using gene
CC therapy and expressing the encoding DNA in vivo, to treat or prevent
CC GPCR-associated disorders, especially in humans. For example, they can
CC be used to treat/prevent cardiomyopathy, atherosclerosis, disorders
CC related to signal processing and metabolic pathway modulation (e.g.
CC obesity, anorexia), diabetes, osteoporosis, Crohn's disease, multiple
CC sclerosis, asthma, cancers, neurodegenerative disorders (e.g. Alzheimer's
CC disease, Parkinson's disorder, Huntington's disease), immune disorders,
CC haematopoietic disorders, developmental diseases, neurological disorders,
CC bacterial, fungal, protozoal and viral infections (e.g. with human
CC immunodeficiency virus (HIV)-1 or HIV-2). They can be used diagnostically
CC to determine the presence of or predisposition to a disease associated
CC with altered levels of the polypeptide in mammals (especially humans) by
CC detecting alterations in polypeptide expression levels relative to
CC control samples. They are useful to identify agents binding polypeptide
CC (e.g. cellular receptors or downstream effectors) and/or agents
CC modulating cellular polypeptide expression or activity, useful as
CC antagonists and agonists in disease treatment.

XX Sequence 1104 BP; 216 A; 318 C; 295 G; 275 T; 0 other;

Query Match 99.7%; Score 1037.8; DB 22; Length 1104;
Best Local Similarity 99.8%; Pred. No. 3.4e-287;
Matches 1039; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGTACAACGGGTCTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATGCCCGC 60

DB 60 ATGTACAACGGGTCTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATGCCCGC 119

QY 61 CTGCTCATTTGGCCCTTTGCTGGGCGCACTAGGCAATGGGTCGCCCTGTGGTTTC 120

DB 120 CTGCTCATTTGGCCCTTTGCTGGGCGCACTAGGCAATGGGTCGCCCTGTGGTTTC 179

SQ	Sequence	888 BP; 174 A; 254 C; 231 G; 229 T; 0 other;
	Query Match	85.3%; Score 888; DB 22; Length 888;
	Best Local Similarity	100.0%; Pred. No. 2.7e-244;
	Matches	888; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	61	CTGCTCATTTGGGCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTGGTTTC 120
Db	1	CTGCTCATTTGGGCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTGGTTTC 60
Qy	121	TGCTTCCACATGAAGACCTGGAAGCCAGCAGCTCTTTACCTTTTCAATTTGGCGTGGCT 180
Db	61	TGCTTCCACATGAAGACCTGGAAGCCAGCAGCTCTTTACCTTTTCAATTTGGCGTGGCT 120
Qy	181	GATTTCTCTCTTATGATCGCTCGCTCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
Db	121	GATTTCTCTCTTATGATCGCTCGCTCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 180
Qy	241	TGGGCTTTTGGGACATTCCTCGCGAGTGGGCTCTTTCAGTTTGGCATGAACAGGGCC 300
Db	181	TGGGCTTTTGGGACATTCCTCGCGAGTGGGCTCTTTCAGTTTGGCATGAACAGGGCC 240
Qy	301	GGGAGCATCGTGTTCCTTACGGTGGTGCCTCGGACAGGTATTTCAAAGTGGTCCACCCC 360
Db	241	GGGAGCATCGTGTTCCTTACGGTGGTGCCTCGGACAGGTATTTCAAAGTGGTCCACCCC 300
Qy	361	CACCACGGGTGAACACTATCTCCACCGGTGGCGCTGGCATCGTCTGCACCTGTGG 420
Db	301	CACCACGGGTGAACACTATCTCCACCGGTGGCGCTGGCATCGTCTGCACCTGTGG 360
Qy	421	GCCTGTGTCCTCGGGAACAGTGTATCTTTTGTGGAGAACCATCTCTCGGTGCAAGAG 480
Db	361	GCCTGTGTCCTCGGGAACAGTGTATCTTTTGTGGAGAACCATCTCTCGGTGCAAGAG 420
Qy	481	ACGGCGTCTCTGTGAGAGCTTATCATGAGTCGAGTCGGCCATGCTGCATGACATATG 540
Db	421	ACGGCGTCTCTGTGAGAGCTTATCATGAGTCGAGTCGGCCATGCTGCATGACATATG 480
Qy	541	TTCCAGCTGGAGTCTTTATCCCTCGGCACTATCTTATTTGCTCTTCAAGATTGT 600
Db	481	TTCCAGCTGGAGTCTTTATCCCTCGGCACTATCTTATTTGCTCTTCAAGATTGT 540
Qy	601	TGGAGCTTGAGGCGGAGCAGCAGCTGCGCCAGACAGGCTCGGATGAAGAGGCCCGG 660
Db	541	TGGAGCTTGAGGCGGAGCAGCAGCTGCGCCAGACAGGCTCGGATGAAGAGGCCCGG 600
Qy	661	TTCATCATGTTGGTGGCAATTTGTTTCATCATGCTACCTGCGCCAGGCTCTGCTAGA 720
Db	601	TTCATCATGTTGGTGGCAATTTGTTTCATCATGCTACCTGCGCCAGGCTCTGCTAGA 660
Qy	721	CTCTATTTCTCTGAGCGGTGCCCTCGAGTGCCTGCGATCCCTCTGTCCATGGGCGCTG 780
Db	661	CTCTATTTCTCTGAGCGGTGCCCTCGAGTGCCTGCGATCCCTCTGTCCATGGGCGCTG 720
Qy	781	CACATAACCTTCAGCTTCACTACATGAACAGCATGCTGGATCCCTCGGTATATTATTT 840
Db	721	CACATAACCTTCAGCTTCACTACATGAACAGCATGCTGGATCCCTCGGTATATTATTT 780
Qy	841	TCAAGCCCTCTCTTCCCAATTTTCAACAAGCTCAAAATCTCAGTCTGAACCCCAAG 900
Db	781	TCAAGCCCTCTCTTCCCAATTTTCAACAAGCTCAAAATCTCAGTCTGAACCCCAAG 840
Qy	901	CAGCCAGGACACTCAAAAACAAAGCCGGAAGAGATGCCAATTTG 948
Db	841	CAGCCAGGACACTCAAAAACAAAGCCGGAAGAGATGCCAATTTG 888
RESULT 11		
AXX16671		
ID	AXX16671	
XX	AXX16671 standard; cDNA; 1361 BP.	
AC	AXX16671;	
XX		

DT	29-APR-1999	(first entry)
XX		
DE	G-protein coupled receptor	HM74A encoding cDNA.
XX		
KW	HM74A receptor; G-protein coupled receptor; infection; pain; cancer;	
KW	diabetes; obesity; neurological disorder; heart failure; hypertension;	
XX	asthma; allergy; ss.	
OS	Homo sapiens.	
XX		
EH	Key	Location/Qualifiers
FT	CDS	61..1152
FT		/*tag= a
FT		/product= "HM74A receptor"
PN	WO9856820-A1.	
XX		
PD	17-DEC-1998.	
XX		
PF	12-JUN-1998;	98WO-US12386.
XX		
PR	12-JUN-1997;	97US-0049480.
XX		
PA	(SMIK) SMITHKLINE BECHAM CORP.	
XX		
PI	Bergsma DJ, Elshourbagy NA, Guerrero SF, Li X, Mooney JL;	
XX		
DR	WPI: 1999-095273/08.	
DR	P-PSDB; AAW94654.	
XX		
PT	New isolated G-protein coupled receptor, HM74A - used to develop	
PT	products for treating e.g. infections, pain, cancers, diabetes,	
PT	obesity, neurological disorders, heart failure, hypertension, asthma	
PT	or allergies	
XX		
PS	Claim 2; Page 30-31; 40pp; English.	
XX		
CC	The present sequence encodes a member of the G-protein coupled receptor	
CC	(7TM receptor) family, designated the HM74A receptor. The proteins,	
CC	agonists, antagonists and polynucleotides can be used for treating	
CC	disorders associated with increased or reduced expression or activity	
CC	of HM74A, e.g. bacterial, fungal, protozoan and viral infections,	
CC	particularly infections caused by HIV-1 or HIV-2, pain, cancers,	
CC	diabetes, obesity, anorexia, bulimia, asthma, Parkinson's disease,	
CC	acute heart failure, hypotension, hypertension, urinary retention,	
CC	osteoporosis, angina pectoris, myocardial infarction, stroke, ulcers,	
CC	asthma, allergies, benign prostatic hypertrophy, migraine, vomiting,	
CC	psychotic and neurological disorders, including anxiety, schizophrenia,	
CC	manic depression, depression, delirium, dementia, and severe mental	
CC	retardation, and dyskinesias such as Huntington's disease or	
CC	Gilles de la Tourette's syndrome. The products can also be used for	
CC	detection, diagnosis and drug screening.	
XX		
SQ	Sequence 1361 BP; 291 A; 390 C; 342 G; 338 T; 0 other;	
	Query Match	35.8%; Score 372.4; DB 20; Length 1361;
	Best Local Similarity	64.2%; Pred. No. 2.1e-96;
	Matches	600; Conservative 0; Mismatches 316; Indels 18; Gaps 2;
Qy	1	ATGTACACGGGTGCTGCTGCCGATCGAGGGGACACCATCTCCAGGTGATGCCGCG 60
Db	97	ATAGACAAGAAGAACTGCTGTGTCTCCGAGATGACTTCATTTGTCGAAGTGTGCCGCG 156
Qy	61	CTGCTCATTTGGGCTTTGTGCTGGCGCACTAGGCAATGGGGTCGCCCTGTGGTTTC 120
Db	157	GTGTTGGGGCTGGAGTTTATCTTCGGGCTTCTGGCAATGGCCTTGCCCTGTGATTTTC 216
Qy	121	TGCTTCCACATGAAGACCTGGAAGCCAGCAGCTCTTTACCTTTTCAATTTGGCGTGGCT 180
Db	217	TGTTTCCACCTCAAGTCTCTGGAATCCAGCGGATTTCTGTTCAACCTGGCAGTGGCT 276
Qy	181	GATTTCTCTCTTATGATCTGCTGCCCTTTTGGGACAGCAATTAACCTCAGACGTAGACAC 240

Db 277 GACCTTTCTACTGATCATCTCCCTGCCCTCTCTGATGGACAACATATGTGAGCGGTGGGAC 336
QY 241 TGGGCTTTTGGGACATCTCCCTGCCAGTGGGCTCTTTCACGTTGGCCATGAACAGGCGC 300
Db 337 TGAAGTTTGGGACATCCCTTGGCGCTGATGCTCTTCTATGTGGCTATGAACCGCCAG 396
QY 301 GGGAGCATCGTGTCCCTTACGGTGGTGGCTGCGGACAGAGTATTTTCAAGTGGTCCACCCC 360
Db 397 GCGAGCATCATCTTCCCTACGGTGGTGGCGGTAGACAGAGTATTTCCGGGTGGTCCATCCC 456
QY 361 CACCACCGGTGAACACTATCTCCACCGCGGTGGCGGCTGGCATCTGCTCCACCCCTGTGG 420
Db 457 CACCACCGCTGACAAAGATCTCCAATCGGACAGAGCCATCTCTTCCCTTCTGTGG 516
QY 421 GCCTGTGTCATCTCCGGAACAGTGTATCTTTTGTCTGGAGAACCATCTCTCTCGTGCAGAG 480
Db 517 GGCATCACTATTGGCGCTGACAGTCCACCTCTCTGAAGAGAGATGCGGATCCAGATGGC 576
QY 481 ACGGCGGTCTCTGTGAGAGCTTCATCATGAGAGTGGGCCAATGGCTGGCATGACATCATG 540
Db 577 GGTGCAAAATTTGTGACAGCAGTTCAGCATCTGCCATACCTTCCAGTGGCAGGAGCCATG 636
QY 541 TTCAGCTGAGTCTTTATGCCCCCTGGGCATCATCTTATTTTGTCTCCCTTCAAGATGTT 600
Db 637 TTCTCTCTGGAGTCTTCTCTCCCTGGGCGATCATCTGTCTGCTCAGCCAGAAATATC 696
QY 601 TGGAGCTGAGCGGAGGAGCAGAGCTGGCCAGACAGGCTCGGATGAAGAGGCGACCCGG 660
Db 697 TGGAGCTGGGCGAGAG---ACAANTGGACCGGATGCCAAGATCAAGAGAGCATCACC 753
QY 661 TTCATCATGTGTGGCAATTTGTTCATCATCATGCTTACCTGCCAGCGTGTCTGTCTAGA 720
Db 754 TTCATCATGTGTGGCATCGTCTTTGTCTCATCTGCTTCTCCAGCGTGGTGTGGCG 813
QY 721 CTCTATTTCTCTGGACGGTGGCCCTC-----GAGTGGCTCGGATCCCTCT 765
Db 814 ATCCGATCTTCTGGCTCTGCACTTCCGGCAGCAGAAATTTGGAAGTGTACCGCTCG 873
QY 766 GTCCATGGGCGCCTGCACATAACCCCTCAGCTTCACCTTACATGAACAGCATGCTGGATCCC 825
Db 874 GTGACCTGGCGTCTTTATCATCTCAGCTTCACCTTACATGAACAGCATGCTGGACCCC 933
QY 826 CTGGTGTATTTTCAAGCCCTCTCTTCCCAAAATTTACAAAGCTCAAAATCTGC 885
Db 934 GTGGTGTACTATTTTCCAGCCCATCTTCCCAACTTCTTCTCCACTTTTGATCAACCCG 993
QY 886 ASCTGTAACCCAGCAGCCAGGACACTCAAAA 919
Db 994 TGGCTCCAGAGGAAGATGACAGGTGAGCCAGATA 1027

RESULT 12

AAS07952
ID AAS07952 standard; cDNA; 1092 BP.

XX AC AAS07952;
XX XX

XX 23-OCT-2001 (first entry)

XX Human cDNA encoding G-protein coupled receptor, hrUP25.

XX Human; G-protein coupled receptor; GPCR; hrUP25; agonist;
KW Inverse agonist; lung cancer; ss.
XX OS Homo sapiens.

XX Key Location/Qualifiers
FH 1..1092
FT CDS /*tag= a
FT /product= "hrUP25"

XX PN W020136471-A2.

PD 25-MAY-2001.
XX 16-NOV-2000; 2000WO-US31509.
XX 17-NOV-1999; 99US-0166088.
PR 17-NOV-1999; 99US-0166099.
PR 17-NOV-1999; 99US-0166369.
PR 23-DEC-1999; 99US-0171900.
PR 23-DEC-1999; 99US-0171901.
PR 23-DEC-1999; 99US-0171902.
PR 11-FEB-2000; 2000US-0181749.
PR 14-MAR-2000; 2000US-0189258.
PR 14-MAR-2000; 2000US-0189259.
PR 10-APR-2000; 2000US-0195898.
PR 10-APR-2000; 2000US-0195899.
PR 10-APR-2000; 2000US-0196078.
PR 28-APR-2000; 2000US-0200419.
PR 12-MAY-2000; 2000US-0203630.
PR 12-JUN-2000; 2000US-0210741.
PR 12-JUN-2000; 2000US-0210982.
PR 21-AUG-2000; 2000US-0226760.
PR 26-SEP-2000; 2000US-0235418.
PR 26-SEP-2000; 2000US-0235779.
PR 20-OCT-2000; 2000US-0242332.
PR 20-OCT-2000; 2000US-0242343.
XX (AREN-) ARENA PHARM INC.
PA Chen R, Dang HT, Lowitz KP;
PI WPI; 2001-355616/37.
XX P-PSDB; AAU04379.
DR Endogenous and non-endogenous versions of human G-protein coupled
XX receptors for direct identification of candidate compounds as agonists,
PT inverse agonists or partial agonists for use as therapeutic agents -
XX Claim 71; Page 120-121; 159pp; English.
XX The sequence encodes a human G-protein coupled receptor (GPCR),
XX hrUP25. The endogenous and non-endogenous, constitutively activated
CC versions of human G-protein coupled receptors (GPCR), are useful for
CC direct identification of candidate compounds as receptor agonists,
CC inverse agonists or partial agonists having applicability as therapeutic
CC agents for treating diseases related to GPCR, e.g. lung cancer.
CC Non-endogenous version of human GPCRs are also utilized in research
CC settings and in vitro and in vivo system, incorporating GPCRs can be
CC utilised to elucidate and understand the roles these receptors
CC play in the human condition, both normal and diseased.
XX SQ Sequence 1092 BP; 223 A; 320 C; 273 G; 276 T; 0 other;

Query Match 35.5%; Score 369.2; DB 22; Length 1092;
Best Local Similarity 64.0%; Pred. No. 1.6e-95;
Matches 598; Conservative 0; Mismatches 318; Indels 18; Gaps 2;

QY 1 ATGTACAAAGGTCGTGCTCGCATCGAGGGGACACCATCTCCAGGTGATCCGCGC 60
Db 37 ATAGACAAGAAGAACTCTGTGTTCGAGATGACTTCAATGTCAAGGTGTTGCCGCGC 96
QY 61 CTGCTCATTTGGCCTTTGTGCTGGCGCACATAGGCAATGGGTCGCCCTGTGTGTTTC 120
Db 97 GTGTGGGGCTGGAGTTTATCTTCGGGCTTCGCGCAATGCGCTTGCCCTGTGGATTTC 156
QY 121 TGCTTCCATGAAAGACCTGGAAGCCAGACCTGTTTACCTTTTCAATTTGGCGGTGGT 180
Db 157 TGTTTCCACCTCAAGTCTCTGGAATCCAGCGGATTTTCTGTTCACACCTGGCAGTGGT 216
QY 181 GATTTCCTCTATGATCTGCTGCTTTTCGGACACACTATTACCTCAGACGTAGACAC 240
Db 217 GACTTTCTGATCATCTGCTGCTGCTTCTCTGATGACAACTATGTGAGCGGTGGGAC 276
QY 241 TGGGCTTTTGGGACATTCCTCTGCCGAGTGGGCTCTTTCACGTTGGCCATGAACAGGCGC 300

	Db	277	TGGAAGTTTGGGACATCCCTTCGCCGGCTGATGCTCTTCATGTGTGCTATGAACCCCGAG	336
	Qy	301	GGAGCATCGTGTTCCCTTAACGGTGGTGGCTGCGGACAGGTATTTCAAAGTGGTGCCACCC	360
	Db	337	GGCAGCATCATCTTCCTCAACGGTGGTGGGGTAGACAGGTATTTCGGGGTGGTGCATCCC	396
	Qy	361	CACCACGGGTGAACACTATCTCCACCCGGGTGGCGGTGGGCATCGCTGCACCTGTGG	420
	Db	397	CACACGCCCTGAACAAGATCTCCAATCGACAGCACGCCATCATCTCTTGCCTTGTGTGG	456
	Qy	421	GCCTGTGTCATCTGGGAACAGGTATCTTTTGCCTGSGAACAACATCTCTGCGTGAAGAAG	480
	Db	457	GGCATCACTATTTGGCCTGACAGTCCACCTCTCTGAAGAAGAAGATCCGATCCAGAAATGGC	516
	Qy	481	ACGGCCGCTCTCCTGTGAGAGCTTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATC	540
	Db	517	GGTGCAATTTTGTGCAGCAGCTTCACGATCTGCCATACCTCCAGTGGCAGCAGGCCATG	576
	Qy	541	TTCCAGCTGGAGTTCTTTATGCCCCTCGGGATCATCTTATTTTGCCTTTCGAAGATGTT	600
	Db	577	TTCTCTCTGGAGTTCTTCTCTGCCCTTGGGCATCATCTCTTCTGCTTCAGCCAGAAATTATC	636
	Qy	601	TGAGAGCTGAGCGGAGGCAGCAGCTGGCCAGACAGGCTCGGATGAAGAAGCGCACCCGG	660
	Db	637	TGAGAGCTTGGGGCAGAG --- ACAAAATGGACGGCGCATGCCAAGATCAAGAGAGCCATCAGC	693
	Qy	661	TTCATCATGGTGTGGCAATTGTGTTTCATCATCATGCTAGCTGCCACAGCGTGCTGCTAGA	720
	Db	694	TTCATCATGGTGTGGCCATCGCTTTTGTCATCTGCTTCTTCCCAGCGTGGTGTGGCG	753
	Qy	721	CTCTATTTCTCTGGACGGTGGCCCTC-----GAGTGCCTTGGCATCCCTCT	765
	Db	754	ATCCGATCTCTTCTGGCTCTGTCACACTTCGGGCACGACAGAATTGTGAAGTGTACCGCTCG	813
	Qy	766	GTCATGGGGCCCTGCACATAACCTTCAGCTTCACCTACATGNACAGCATGCTGGATCCC	825
	Db	814	GTGGACCTGGCGGTTCTTTATCATCTCTCAGCTTCACCTTACATGAACAGCATGTGGACCCC	873
	Qy	826	CTGGTGATTTATTTTCAAGCCCTCTTTCCCAAATCTTACAACAAGCTCAAAATCTCG	885
	Db	874	GTGGTGTACTIONTTCTCCAGSCCATCTTTTCCCAACTTCTTCTCCACTTTGATCAACCGC	933
	Qy	886	AGTCTGMAACCCCAAGCAGCCAGCACTCAAAA	919
	Db	934	TGCTCTCAGAGGAAGATGACAGGTGAGCCAGATA	967

RESULT 13
AAA30658
ID AAA30658 standard; cDNA; 1164 BP.

Behan DP, Chalmers DT, Liaw CW;
WPI; 2000-329165/28.
P-PSDB; AAY90637.

Non-endogenous constitutively activated human G protein-coupled receptors, useful for identifying agonists for use as pharmaceutical agents -

Example 1; Page 185; 341pp; English.

The invention relates to constitutively active, non-endogenous versions of endogenous human orphan G protein-coupled receptors (GPCRs, AAY90643- AAY90677 and AAY90683-Y90687), and to DNA encoding them (AAA30709-A30743 and AAA30775-A30779). The mutant proteins of the invention contain a mutation in a portion of the protein comprising intracellular loop 3 (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X, is substituted for an endogenous residue in IC3 at a position 16 amino acids N-terminal of an endogenous proline in TM6 to form a sequence X-ALA15-Pro. The endogenous amino acid is selected from Lys, His, Arg or Ala, and is preferably Lys. When the endogenous residue at this position is Lys, this residue is replaced by His, Arg or preferably Ala. The 15 amino acid stretch between the substituted amino acid and the Pro may be endogenous, non-endogenous, or a mixture of endogenous and non-endogenous residues. The constitutively active GPCRs are useful for identifying antagonists, agonists and partial agonists for use as pharmaceutical agents. The mutant proteins are also useful in research settings for elucidating the roles of the receptors in normal and diseased conditions. Antagonists for a particular GPCR are useful for treating diseases and disorders associated with that receptor. Because the novel mutant GPCRs are constitutively active, they can be used directly for screening of compounds without the need for endogenous ligands. The present sequence represents cDNA encoding a human wild-type GPCR used in an exemplification of the invention. This was cloned and subjected to site-directed mutagenesis (SDM) to generate DNA encoding the corresponding mutant of the invention.

Sequence 1164 BP; 246 A; 337 C; 288 G; 293 T; 0 other;

Db 457 GGCATCACTGTTGGCCCTAACAGTCCACCTCTGAAGAAGAAGTTGCTGATCCAGAATGGC 516
Qy 481 ACGGCCGTCCTCTGTGAGAGCTTCATCATGGAGTGGCCGAATGGCTGGCATGACATCATG 540
Db 517 CCGCAAAATGTGTCATCAGCTTCAGCATCTGCCATACCTTCCGGTGGCAGCAAGCTATG 576
Qy 541 TTCAGTGGAGTCTTTATGCCCCCTCGGCATCATCTTATTTGCTCCCTTCAAGATTGTT 600
Db 577 TTCCTCTGGAGTTCCTCTGCCCCCTGGGCATCATCTGTTCTGCTCAGCCAGATATC 636
Qy 601 TGGAGCTGAGCGGAGGAGCAGCTGAGCTGGCCAGACAGCTCGGATGATGAAGAGCGACCCGG 660
Db 637 TGGAGCTGCGGCAGAG---ACAAATGGACCGCATGCCAAGATCAAGAGAGCATCACC 693
Qy 661 TTCATCATGCTGTCGCAATGTGTTTATCATCATGCTACCTGCTACCTGCCAGCGTGTCTGTAGA 720
Db 694 TTCATCATGCTGTCGCAATGTGTTTATCATCATGCTACCTGCTACCTGCCAGCGTGTCTGTAGA 753
Qy 721 CTCATATTCCTGCGAGGGTGGCCCTC-----GAGTGGCTCGCATCCCTCT 765
Db 754 ATCCGATCTTCCTGGCTCTGCTGACACTTCGGGCACGCGAAGTTGTAAGTGTACCGCTCG 813
Qy 766 GTCCATGGGCGCTGCACATAACCTCAGCTTACCTTACATGATGACAGCATGCTGGATCCC 825
Db 814 GTGGACCTGGCGTCTTTATCACTCTCAGCTTACCTTACATGATGACAGCATGCTGGACCCC 873
Qy 826 CTGGTGTATTTTCAAGCCCTCTCTTCCCAAAATTTTCAACAGCTCAAAATCTGC 885
Db 874 GTGGTGTACTACTTCTCAGCCCATCTCTTCCCAAAATTTTCAACAGCTCAAAATCTGC 933
Qy 886 AGTCTGAACCCCAAGCAGCCAGGACACTCAAAA 919
Db 934 TGCCTCAGAGGAAGATGACAGGTGAGCCAGATA 967

RESULT 14

ID AAA30738 standard; DNA: 1164 BP.
XX AAA30738;
AC
XX
XX
DT 21-AUG-2000 (first entry)
XX
DE DNA encoding human mutant G protein-coupled receptor HM74 (I230K).
XX
XX G protein-coupled receptor; GPCR; constitutively active;
KW Intracellular loop 3; transmembrane domain 6; drug screening;
KW agonist; antagonist; mutant; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX WO20002129-A1.
PN
XX
XX 20-APR-2000.
PD
XX
XX 12-OCT-1999; 99WO-US23938.
PE
XX
XX 13-OCT-1998; 98US-0170496.
PR
XX
XX (AREN-) ARENA PHARM INC.
XX
XX Behan DP, Chalmers DT, Liaw CW;
PI
XX
XX WPI; 2000-329165/28.
DR P-PSDB; AAY90672.
XX
XX Non-endogenous constitutively activated human G protein-coupled
PT receptors, useful for identifying agonists for use as pharmaceutical
PT agents -
XX
XX Example 2; Page 285-286; 341pp; English.
PS

XX
CC The invention relates to constitutively active, non-endogenous versions
CC of endogenous human orphan G protein-coupled receptors (GPCRs, AAY90643-
CC AAY90677 and AAY90683-Y90687), and to DNA encoding them (AAA30709-A30743
CC and AAA30775-A30779). The mutant proteins of the invention contain a
CC mutation in a portion of the protein comprising intracellular loop 3
CC (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X,
CC is substituted for an endogenous residue in IC3 at a position 16 amino
CC acids N-terminal of an endogenous proline in TM6 to form a sequence
CC X-(AA)15-Pro. The endogenous amino acid is selected from Lys, His, Arg
CC or Ala, and is preferably Lys. When the endogenous residue at this
CC position is Lys, this residue is replaced by His, Arg or preferably Ala.
CC The 15 amino acid stretch between the substituted amino acid and the pro
CC may be endogenous, non-endogenous, or a mixture of endogenous and
CC non-endogenous residues. The constitutively active GPCRs are useful for
CC identifying antagonists, agonists and partial agonists for use as
CC pharmaceutical agents. The mutant proteins are also useful in research
CC settings for elucidating the roles of the receptors in normal and
CC diseased conditions. Antagonists for a particular GPCR are useful for
CC treating diseases and disorders associated with that receptor. Because
CC the novel mutant GPCRs are constitutively active, they can be used
CC directly for screening of compounds without the need for endogenous
CC ligands. Sequences AAA30709-AAA30743 and AAA30775-A30779 represent DNAs
CC encoding the mutant human GPCRs of the invention.
XX
SQ Sequence 1164 BP; 248 A; 335 C; 289 G; 292 T; 0 other;

Query Match 35.2%; Score 366; DB 21; Length 1164;

Best Local Similarity 63.8%; Pred. No. 1.3e-94;

Matches 596; Conservative 0; Mismatches 320; Indels 18; Gaps 2;

Qy 1 ATGTACAAACGGGTGCTGCTGCCGATCGAGGGGACACACCATCTCCAGGTGATCCCGCG 60
Db 37 ATAGACAAGAAGAACTGCTGTGTTCCGAGATGACTTCATTCGCAAGGTGTTGCCCGCG 96
Qy 61 CTGCTCATTTGGCCTTTGTGCTGGCGCCTAGGCAATGGGCTGCGCCCTGTGTGGTTTC 120
Db 97 GTGTTGGGCTGGAGTTTATCTTTGGGCTTCTGGCAATGGCCTTGCCCTGTGGATTTC 156
Qy 121 TGCTTCCACATGAGACCTGGAAGCCAGCAGCTGTTTACCTTTTCAATTTGGCGTGGCT 180
Db 157 TGTTCCACCTCAAGTCTTGGAAATCCAGCCGGATTTTCTCTCAACCTGGCAGTAGT 216
Qy 181 GATTTCTCCTTATGATCTGCTGCTTTCGGGACAGACTATTACCTCAGACGTAGACAC 240
Db 217 GACTTCTACTGATCATCTGCTGCTGCTGATGACTACTATGTGCGCGCTTACAGAC 276
Qy 241 TGGGCTTTTGGGACATTCCTCGCGAGTGGGCTCTTCACGTTGGCCATGAACAGGGCC 300
Db 277 TGGAAATTTGGGACATCCCTTGCCGCTGCTGCTTCTATGTTTGGCATGAACGCCAG 336
Qy 301 GGGAGCATGTTTCTTACGTTGCTGCTGCGGACAGGATTTTCAAGTGGTCCACCCC 360
Db 337 GGCAGCATCATCTTCTCAGCGGTGGGTAGACAGGTATTTCCGGGTGGTCCATCCC 396
Qy 361 CACCACGCGGTGAACACTATCTCCACCGGGTGGCGCTGGCATCTGCTGCAACCTGTGG 420
Db 397 CACCACGCGGTGAACAGATCTCAATTTGACACAGCCATCATCTCTGCTTCTGTGG 456
Qy 421 GCCTGTGTCATCTCTGGGAACAGTGTATCTTTTGTGGGAACCAATCTCTGCGGTCAAGAG 480
Db 457 GGCATCACTGTTGGCCTTAACAGTCCACCTCTCTGAAGAAGAAGTTGCTGATCCAGAATGGC 516
Qy 481 ACGGCCGTCCTCTGTGAGAGCTTCATCATGGAGTCGGCCCAATGGCTGGCATGACATCG 540
Db 517 CCTGCAAAATGTGTGCATCAGCTTTCAGCATCTGCCATACCTTCCGGTGGCAGAGTATG 576
Qy 541 TTCAGCTGGAGTCTTTATGCCCTCGGCATCATCTTATTTTCTCTCTTCAAGATTGTT 600
Db 577 TTCCTCTGGAGTCTCTCTGCCCTCGGCATCATCTGTTCTGCTCAGCCAGATATTC 636
Qy 601 TGGAGCCTGAGCGGAGGAGCAGCTGCGCCAGACAGCTCGGATGAAGAAGCGACCCCG 660

Db 637 TGGAGCCTGGCGCAGAG---ACAATGACCGGCATGCCAAGATCAAGAGAGCAAAACC 693
Qy 661 TTCATCATGGTGGGCAATGTGTTTCATCATGCTACCTGCGCCAGCGTGTCTGCTAGA 720
Db 694 TTCATCATGGTGGGCAATGTGTTTCATCATGCTACCTGCGCCAGCGTGTGTCGGG 753
Qy 721 CTCATTTTCCTCTGGACGGTCCCTC-----GAGTGCCTGGGATCCCTCT 765
Db 754 ATCCGACCTGCGGCTCTTTCATCTGACACTTCGGGCGCAGCAATTTGGAAGTGTACCGCTCG 813
Qy 766 GTCCATGGGGCCCTGCACATACCTGCTACCTTACCTACATGACAGCATGCTGGATCCC 825
Db 814 GTGACCTGGCGCTCTTTCATCTGACACTTCAGCTTACCTTACATGACAGCATGCTGGACCCC 873
Qy 836 CTGTGTATTTATTTTCAAGCCCTCTTTCCTCCAAATTTCTACACAGCTCAAAATCTGC 885
Db 874 GTGTGTACTTCTCTCAGCCCATCTTTCCTCCAACTTCTTCTCCATCTTCTCCATCTTATCAACCGC 933
Qy 886 AGTCTGAAACCAAGCAGCCAGGACACTCAAAAA 919
Db 934 TGCCTCCAGGAGAGATGACAGGTGAGCCAGATA 967
RESULT 15
AAQ55054
ID AAQ55054 standard; DNA; 546 BP.
XX AC AAQ55054;
XX AC AAQ55054;
DT 18-JUL-1994 (first entry)
XX DE Sequence of orphan receptor DNA fragment PCR 9.
XX KW PCR; oligo; primer; peptide ligand receptor; ss.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT misc_feature 1..26
FT /*tag= a
FT /label= PCR oligo no. 8
FT misc_feature 526..545
FT /*tag= b
FT /label= PCR oligo no. 10
XX EP578962-A.
XX 19-JAN-1994.
XX 04-JUN-1993; 93EP-0108984.
XX 17-JUL-1992; 92US-0915966.
XX (AMCY) AMERICAN CYANAMID CO.
XX Hadcock JR, Ozenberger BA, Pausch MH;
XX WPI; 1994-017562/03.
XX Isolation and identifying new receptor DNA - pref. G-protein
XX linked receptors, using PCR with new receptor specific primers
XX Claim 6; Fig 3; 17pp; English.
XX Consensus sequences are identified from nine rat gene receptors. The
XX receptors are for somatostatin, substance K, substance P, neuromedin
XX K, thyrotropin, HH/CG and others designated mas, mrg and rta. Four
XX regions of nucleotide sequence are found to exhibit a degree of
XX conservation significant enough to design degenerate oligo primers
XX for PCR (see AAQ55045-49). Five oligos are designed with 4- to 64-fold
XX direct degeneracy plus 5'-23' inosines. The oligos are designed to
XX anneal sequences in G protein-linked peptide receptor genes. They
XX are used as six different primer pairs for PCR using rat genomic DNA

CC as template. All six reactions produce DNA fragments of the expected
CC size. DNA fragments are cloned and two samples, designated PCR9 and
CC PCR11(AAQ55054,055055), are found to contain sequences characteristic
CC of G protein-linked receptors. Each fragment contains a single
CC ORF. There is considerable similarity between rat somatostatin
CC receptor and PCR11, suggesting that PCR11 is a novel somatostatin
CC receptor subtype. Both fragments exhibit not only primary sequence
CC similarity to other receptors but also the hydrophathy pattern
CC characteristic of G protein-linked receptors. The novel G protein-
CC linked receptors are possibly of the peptide ligand subclass.
XX
SQ Sequence 546 BP; 106 A; 157 C; 144 G; 139 T; 0 other;
Query Match 21.2%; Score 220.2; DB 15; Length 546;
Best Local Similarity 65.0%; Pred. No. 5.9e-53;
Matches 357; Conservative 0; Mismatches 188; Indels 4; Gaps 2;
Qy 157 TACCTTTTCAATTTGGCGTGGCTGA-TTTCCTCCTTATGATCTGCTGCTTTTCGGAC 215
Db 1 TTCGTGGTGAACCTGGTGGGCTGACTTTTCTCCTGATCATTTGCTTGGCTTCTTTCGAC 60
Qy 216 AGACTATTACCTCAGACGTAGACACTGGGCTTTTGGGGACATTCCTGCCGAGTGGGCT 275
Db 61 GGACAACATATCTCCAGAACTGGGACTGGAGGTTCGGGAGCATCCCTCCCGTGATGCT 120
Qy 276 CTTACGTTTGGCCATGAACAGGCGCGGAGCATGCTGTTCCTTACGGTGGTGGCTGCGGA 335
Db 121 CTTTCATGTTGGCCATGAACAGGCGGAGCATCATCTTCCTCAGGTGGTGGCTGTGA 180
Qy 336 CAGGTATTTCAAGTGTGCTCCACCCCGGCGGTGAACACTATCTCCACCCCGGCTGGC 395
Db 181 CAGGTACTTTCAGGTGGTGGTCCACCCCGGCGGCTTCTTCTTGAACAGATCTCCAACCGGACGC 240
Qy 396 GGCTGGCATCTGTCACCCCTGTGGCCCTGCTGATCTCTGGGAACAGTATCTTTTCT 455
Db 241 GGCAATCATCTTCTTGTCTTCTTGGGCGATCACCATCGGCTGACAGTCCACCTCCTCTA 300
Qy 456 GGAGAACCATCTCTGCGTGAAGAGAGACGGCGCTCTCTCTGTGAGAGCTTTCATCATGGAGTC 515
Db 301 CACGACATGATGACCCGAAACGGCGATGCAAACTGTGACAGCATTTTAGCATCTGCTA 360
Qy 516 GGCAATGGCTGGCATGACATCATGTTCCAGCTGGAGTCTTTTATGCCCTCCGGCATCAT 575
Db 361 CACTTTCAGGTGGCAGCATGCAATGTTCTCTTGAATTTCTCTGCTGCTGCTGGCATCAT 420
Qy 576 CTTATTTTGTCTCTCAAGATTGTTTGGAGCTGAGGCGGAGGAGCAGCTGGCCAGACA 635
Db 421 CCTGTCTCTGCTCTGGCAGGATCATTTGGAGCCTAAGGCAGAG---ACAGATGGACAGCA 477
Qy 636 GGCTCGGATGAAGAGCGACCCGCTTTCATCATGCTGGTGGCAATTTGTGTTCATCATCATG 695
Db 478 CGTCAAGATCAAGAGGCGCCATCACTTCATCATGCTGGTGGCTGCCATTTGTTGGCATCTG 537
Qy 696 CTACCTGCC 704
Db 538 CTGGCTGCC 546

Search completed: October 29, 2002, 02:41:38
Job time : 220 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2002, 03:44:28 ; Search time 61 seconds
(without alignments)
630.025 Million cell updates/sec

Title: US-09-886-041-2

Perfect score: 1853

Sequence: 1 MYNGSCRRIGDTISQVMP.....ANSFQSGDQWDPHIVEWH 346

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1853	100.0	346	22	ABB44522 Human GPCR1a polyp
2	1853	100.0	346	22	AAU06197 Novel human G prot
3	1853	100.0	346	22	AAU04373 Human G-protein co
4	1853	100.0	346	22	AAU04373 Human G-protein co
5	1853	99.2	346	22	AAU04373 Human GPCR11 #2.
6	1570	84.7	296	22	AAU04373 Human GPCR11 #1.
7	883.5	47.7	363	20	AAU04373 G-protein coupled
8	883.5	47.7	363	20	AAU04373 Human G-protein co
9	881.5	47.6	387	21	AAU06072 Human mutant G pro
10	880.5	47.5	387	21	AAU06072 Human G protein-co
11	529	28.5	384	22	AAU02493 Human CON103 G pro

12	529	28.5	423	20	AAW88460 Human 7-transmembr
13	529	28.5	423	22	AAU04365 Human oploid-type
14	529	28.5	423	22	AAU04365 Human G-protein co
15	529	28.5	455	21	AAU04365 Human cell surface
16	510	27.5	476	20	AAW88461 Human 7-transmembr
17	451.5	24.4	319	21	AAU04365 Human G-protein co
18	448.5	24.2	319	21	AAU04365 Human G-protein co
19	448.5	24.2	319	21	AAU04365 Human mutant G pro
20	448.5	24.2	319	22	AAU04365 Human G-protein co
21	366.5	19.8	341	22	AAU04365 Human G-protein co
22	366.5	19.8	341	22	AAU04365 Human G-protein co
23	366.5	19.8	346	22	AAU04365 Human P2Y-like GPC
24	366.5	19.8	346	22	AAU04365 Human G-protein co
25	366.5	19.8	346	22	AAU04365 Human G-protein co
26	366.5	19.8	346	22	AAU04365 Human G-protein co
27	366.5	19.8	346	22	AAU04365 Human G-protein co
28	366.5	19.8	346	22	AAU04365 Human G-protein co
29	362	19.5	345	22	AAU04365 Human G-protein co
30	361.5	19.5	346	22	AAU04365 Human G-protein co
31	357	19.3	331	22	AAU04365 Human G-protein co
32	357	19.3	347	22	AAU04365 Human G-protein co
33	355.5	19.2	309	22	AAU04365 Human G-protein co
34	355.5	19.2	309	22	AAU04365 Human G-protein co
35	354	19.1	373	22	AAU04365 Human G-protein co
36	346.5	18.7	370	19	AAU04365 Human G-protein co
37	336	18.1	339	15	AAU04365 Human G-protein co
38	336	18.1	339	18	AAU04365 Human G-protein co
39	336	18.1	339	19	AAU04365 Human G-protein co
40	336	18.1	339	21	AAU04365 Human G-protein co
41	336	18.1	339	21	AAU04365 Human G-protein co
42	336	18.1	339	22	AAU04365 Human G-protein co
43	335	18.1	339	21	AAU04365 Human G-protein co
44	327	17.6	309	21	AAU04365 Human G-protein co
45	327	17.6	309	21	AAU04365 Human G-protein co

ALIGNMENTS

RESULT 1

ABB44522
ID ABB44522 standard; Protein; 346 AA.

AC ABB44522;

DT 28-JAN-2002 (first entry)

XX Human GPCR1a polypeptide SEQ ID NO 2.

XX Human; GPCR; G-coupled protein-receptor; cardiant; antiarteriosclerotic;
XX anabolic; cytotstatic; antiviral; gene therapy; cardiomyopathy; obesity;
XX anorexia; diabetes; osteoporosis; Crohn's disease; multiple sclerosis;
XX asthma; Alzheimer's disease; Parkinson's disorder; Huntington's disease;
XX infection; human immunodeficiency virus; HIV.

OS Homo sapiens.

PN WO200174904-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US10241.

XX 31-MAR-2000; 2000US-193664P.

XX 05-APR-2000; 2000US-194614P.

XX 06-APR-2000; 2000US-195063P.

XX 06-APR-2000; 2000US-195066P.

XX 06-APR-2000; 2000US-195067P.

XX 06-APR-2000; 2000US-195068P.

XX 06-APR-2000; 2000US-195069P.

XX 06-APR-2000; 2000US-195070P.

XX 21-JUL-2000; 2000US-219855P.

Query Match 100.0%; Score 1853; DB 22; Length 346;
Best Local Similarity 100.0%; Pred. No. 1.2e-199;
Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYNGSCRIEGDTISQVMPPLLIIVAFVLGALNGVALCGFCFHMKTWKPTVYLFNLAVA 60
DB 1 MYNGSCRIEGDTISQVMPPLLIIVAFVLGALNGVALCGFCFHMKTWKPTVYLFNLAVA 60
QY 61 DFLMTCCLPRTDYLLRRRHAFGDIPCRVGLFTLAMNRAGSIVELTVVAADRYFKVHP 120
DB 61 DFLMTCCLPRTDYLLRRRHAFGDIPCRVGLFTLAMNRAGSIVELTVVAADRYFKVHP 120
QY 121 HHAVENTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCSFIMESANGWHDIM 180
DB 121 HHAVENTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCSFIMESANGWHDIM 180
QY 181 FOLEFFMPLGIILFCFVKIWSLRRRQQLARQARMKATRFIMVAIVFTCYLPSVSAR 240
DB 181 FOLEFFMPLGIILFCFVKIWSLRRRQQLARQARMKATRFIMVAIVFTCYLPSVSAR 240
QY 241 LYFLWTVPPSACDPSVHGALHITLSFTYMNMLDPLVYFSSFPKFKYKIKCSLKPK 300
DB 241 LYFLWTVPPSACDPSVHGALHITLSFTYMNMLDPLVYFSSFPKFKYKIKCSLKPK 300
QY 301 QPGHKTQRPPEEMPISNLGRSCISVANSFQSDGQWDPHIVEWH 346
DB 301 QPGHKTQRPPEEMPISNLGRSCISVANSFQSDGQWDPHIVEWH 346

RESULT 3
AAU04373
ID AAU04373 standard; Protein; 346 AA.
AC AAU04373;
XX
XX
DT 23-OCT-2001 (first entry)
DE Human G-protein coupled receptor, hRUP19.
XX
XX
KW Human: G-protein coupled receptor; GPCR; hRUP19; agonist;
KW Inverse agonist; lung cancer.
OS
XX
XX
PN Homo sapiens.
XX
XX
PD WO200136471-A2.
XX
PF 25-MAY-2001.
XX
XX
PR 16-NOV-2000; 2000WO-US31509.
XX
PR 17-NOV-1999; 99US-0166088.
PR 17-NOV-1999; 99US-0166099.
PR 17-NOV-1999; 99US-0166369.
PR 23-DEC-1999; 99US-0171900.
PR 23-DEC-1999; 99US-0171901.
PR 23-DEC-1999; 99US-0171902.
PR 11-FEB-2000; 2000US-0181749.
PR 14-MAR-2000; 2000US-0189258.
PR 14-MAR-2000; 2000US-0203630.
PR 12-MAY-2000; 2000US-0203630.
PR 12-JUN-2000; 2000US-0210741.
PR 12-JUN-2000; 2000US-0210982.
PR 21-AUG-2000; 2000US-0226760.
PR 26-SEP-2000; 2000US-0235418.
PR 26-SEP-2000; 2000US-0235779.
PR 20-OCT-2000; 2000US-0242332.
PR 20-OCT-2000; 2000US-0242343.
XX

(AREN-) ARENA PHARM INC.

Chen R, Dang HT, Lowitz KP;

WPI; 2001-355616/37.

DR N-PSDB; AAS07946.

XX Endogenous and non-endogenous versions of human G-protein coupled
PT receptors for direct identification of candidate compounds as agonists,
PT inverse agonists or partial agonists for use as therapeutic agents -

PS Claim 45; Page 110-111; 160pp; English.

XX The sequence represents a human G-protein coupled receptor (GPCR),
CC hRUP19. The endogenous and non-endogenous, constitutively activated
CC versions of human G-protein coupled receptors (GPCR), are useful for
CC direct identification of candidate compounds as receptor agonists,
CC inverse agonists or partial agonists having applicability as therapeutic
CC agents for treating diseases related to GPCR, e.g. lung cancer.
CC Non-endogenous version of human GPCRs are also utilized in research
CC settings and in vitro and in vivo system, incorporating GPCRs can be
CC utilised to elucidate and understand the roles these receptors
CC play in the human condition, both normal and diseased.

SQ Sequence 346 AA;

Query Match 100.0%; Score 1853; DB 22; Length 346;

Best Local Similarity 100.0%; Pred. No. 1.2e-199; Indels 0; Gaps 0;
Matches 346; Conservative 0; Mismatches 0;

QY 1 MYNGSCRIEGDTISQVMPPLLIIVAFVLGALNGVALCGFCFHMKTWKPTVYLFNLAVA 60

DB 1 MYNGSCRIEGDTISQVMPPLLIIVAFVLGALNGVALCGFCFHMKTWKPTVYLFNLAVA 60

QY 61 DFLMTCCLPRTDYLLRRRHAFGDIPCRVGLFTLAMNRAGSIVELTVVAADRYFKVHP 120

DB 61 DFLMTCCLPRTDYLLRRRHAFGDIPCRVGLFTLAMNRAGSIVELTVVAADRYFKVHP 120

QY 121 HHAVENTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCSFIMESANGWHDIM 180

DB 121 HHAVENTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCSFIMESANGWHDIM 180

QY 181 FOLEFFMPLGIILFCFVKIWSLRRRQQLARQARMKATRFIMVAIVFTCYLPSVSAR 240

DB 181 FOLEFFMPLGIILFCFVKIWSLRRRQQLARQARMKATRFIMVAIVFTCYLPSVSAR 240

QY 241 LYFLWTVPPSACDPSVHGALHITLSFTYMNMLDPLVYFSSFPKFKYKIKCSLKPK 300

DB 241 LYFLWTVPPSACDPSVHGALHITLSFTYMNMLDPLVYFSSFPKFKYKIKCSLKPK 300

QY 301 QPGHKTQRPPEEMPISNLGRSCISVANSFQSDGQWDPHIVEWH 346

DB 301 QPGHKTQRPPEEMPISNLGRSCISVANSFQSDGQWDPHIVEWH 346

RESULT 4

AAG80968

ID AAG80968 standard; Protein; 346 AA.

XX AAG80968;

XX AAG80968;

DT 28-AUG-2001 (first entry)

XX Human nGPCR11 #2.

DE Human nGPCR11 #2.

XX

KW G protein-coupled receptor; nGPCR; seven transmembrane receptor;

KW signal transduction; schizophrenia; thyroid disorder; renal failure;

KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;

KW cardiovascular disease; proliferative disorder; hormonal disorder;

KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;

KW attention deficit-hyperactivity disorder/attention deficit disorder;

KW Parkinson's disease; migraine; senile dementia; inflammatory disease;

KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;

PT treatment or prevention of GPCR associated disorders e.g.
PT cardiomyopathy or atherosclerosis, and to screen for antagonists and
PT agonists useful therapeutically
XX Claim 1; Page 11; 157pp; English.
XX The invention relates to nucleic acid sequences (ABA81529-ABA81552) that
CC encode G-coupled protein-receptor related polypeptides
CC (ABA84522-ABA84543). The isolated polypeptide having a sequence differing
CC by no more than 15 % of amino acid residues from one of 22 amino acid
CC sequences (or mature forms of the sequences), fully defined in the
CC specification and corresponding to human G-protein coupled receptor x
CC (GPCRx) polypeptides. The polypeptides have potential cardiant,
CC antiarteriosclerotic, anabolic, cytostatic and antiviral activity. The
CC polypeptides can be administered therapeutically, especially using gene
CC therapy and expressing the encoding DNA in vivo, to treat or prevent
CC GPCR-associated disorders, especially in humans. For example, they can
CC be used to treat/prevent cardiomyopathy, atherosclerosis, disorders
CC related to signal processing and metabolic pathway modulation (e.g.
CC obesity, anorexia), diabetes, osteoporosis, Crohn's disease, multiple
CC sclerosis, asthma, cancers, neurodegenerative disorders (e.g. Alzheimer's
CC disease, Parkinson's disorder, Huntington's disease), immune disorders,
CC haematopoietic disorders, developmental diseases, neurological disorders,
CC bacterial, fungal, protozoal and viral infections (e.g. with human
CC immunodeficiency virus (HIV)-1 or HIV-2). They can be used diagnostically
CC to determine the presence of or predisposition to a disease associated
CC with altered levels of the polypeptide in mammals (especially humans) by
CC detecting alterations in polypeptide expression levels relative to
CC control samples. They are useful to identify agents binding polypeptide
CC (e.g. cellular receptors or downstream effectors) and/or agents
CC modulating cellular polypeptide expression or activity, useful as
CC antagonists and agonists in disease treatment.
XX Sequence 346 AA;

Query Match 99.2%; Score 1839; DB 22; Length 346;
Best Local Similarity 99.4%; Pred. No. 4.7e-198;
Matches 344; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MYNSCCRIEGDTISQVMPPLLIIVAFVLGALNGVALCGCFHMKTWKSTVYLFNLAVA 60
Db 1 MYNSCCRIEGDTISQVMPPLLIIVAFVLGALNGVALCGCFHMKTWKSTVYLFNLAVA 60
QY 61 DFLMLICLPRTDYLLRRHWAEGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVWHP 120
Db 61 DFLMLICLPRTDYLLRRHWAEGDIPCRVGLFTLAMNRAGSIVFLTVVAAGRYFKVWHP 120
QY 121 HHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCSFIMESANGWHDIM 180
Db 121 HHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCSFIMESANGWHDIM 180
QY 181 FQLEFFMPLGIILFCSPKIVWSLRRQOLARQARMKKATRFIMVAIVFTICVLPVSAR 240
Db 181 FQLEFFMPLGIILFCSPKIVWSLRRQOLARQARMKKATRFIMVAIVFTICVLPVSAR 240
QY 241 LYFLWTPVPSACDPSVHGALHITLSFTYMSMDPLVYVYSSPFPKFKYKLIKCSLKPK 300
Db 241 LYFLWTPVPSACDPSVHGALHITLSFTYMSMDPLVYVYSSPFPKFKYKLIKCSLKPK 300
QY 301 QPGHSTQRPPEMPISNLGRRCISVANSFQSDGOWDPHIVEWH 346
Db 301 QPGHSTQRPPEMPISNLGRRCISVANSFQSDGOWDPHIVEWH 346

RESULT 6
AAG80934

ID AAG80934 standard; Protein: 296 AA.

XX AAG80934;

XX AAG80934;

DT 28-AUG-2001 (first entry)

XX Human nGPCR11 #1.

XX G protein-coupled receptor; nGPCR; seven transmembrane receptor;
KW signal transduction; schizophrenia; thyroid disorder; renal failure;
KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
KW cardiovascular disease; proliferative disorder; hormonal disorder;
KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
KW attention deficit-hyperactivity disorder/attention deficit disorder;
KW Parkinson's disease; migraine; senile dementia; inflammatory disease;
KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;
KW neuroprotective.
XX Homo sapiens.
OS WO200136473-A2.
PN 25-MAY-2001.
XX 16-NOV-2000; 2000WO-US31581.
XX 16-NOV-1999; 99US-0165838.
PR 17-NOV-1999; 99US-0166071.
PR 19-NOV-1999; 99US-0166678.
PR 28-DEC-1999; 99US-0173396.
PR 22-FEB-2000; 2000US-0184129.
PR 28-FEB-2000; 2000US-0185421.
PR 28-FEB-2000; 2000US-0185554.
PR 02-MAR-2000; 2000US-0186530.
PR 03-MAR-2000; 2000US-0186811.
PR 09-MAR-2000; 2000US-0188114.
PR 17-MAR-2000; 2000US-0190310.
PR 21-MAR-2000; 2000US-0190800.
PR 20-APR-2000; 2000US-0198568.
PR 02-MAY-2000; 2000US-0201190.
PR 08-MAY-2000; 2000US-0203111.
PR 25-MAY-2000; 2000US-0207094.
XX (PHAA) PHARMACIA & UPJOHN CO.
PA Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;
PI Schellin KA, Kayces PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;
XX WPI; 2001-389826/41.
DR N-PSDB; AAH50974.
XX New G protein-coupled receptor (nGPCR-x) and its encoding
PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -
XX Claim 37; Pages 77-78; 261pp; English.

PS The present invention relates to novel G protein-coupled receptors
XX (nGPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
CC 28, 31-38, 40, 41, 53-60) and their coding sequences. The present
CC sequence is one such G protein-coupled receptor. GPCRs are also known as
CC seven transmembrane receptors and function in signal transduction. The
CC nGPCRx coding sequences are useful for screening a human to diagnose a
CC disorder affecting the brain or a genetic predisposition, specifically
CC schizophrenia. nGPCRx are useful for identifying compounds useful for
CC treating schizophrenia. Detection of nGPCRx in a sample is useful as a
CC diagnostic tool for diseases or disorders e.g. thyroid disorders, renal
CC failure, rheumatoid arthritis, CNS disorders, infections such as HIV-1,
CC metabolic and cardiovascular diseases, proliferative disorders and
CC hormonal disorders. Modulators of nGPCRx activity have the utility for
CC treating neurological disorders, including schizophrenia, ADHD/ADD
CC (attention deficit-hyperactivity disorder/attention deficit disorder),
CC and neuronal disorders such as Alzheimer's disease, Parkinson's disease,
CC migraine and senile dementia. Additional disorders include inflammatory
CC conditions (e.g. Crohn's disease), rheumatoid arthritis, autoimmune
CC disorders, cancers, respiratory ailments such as asthma, and inflammatory
CC diseases e.g. inflammatory bowel disease.

XX Sequence 296 AA;

Query Match 84.7%; Score 1570; DB 22; Length 296;

Best Local Similarity 100.0%; Pred. No. 6.8e-168; Matches 296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LLIVAFVGLGNGVLCGFCGFMKTPSTVYLFNLAADFLMCLPFTDYLLRRH 80
Db 1 LLIVAFVGLGNGVLCGFCGFMKTPSTVYLFNLAADFLMCLPFTDYLLRRH 60

Qy 81 WAFGDIICRVGLFTLNMNAGSIVFLTVAAADRYFKVVPVPHAVNTISTRVAAGIVCTLW 140
Db 61 WAFGDIICRVGLFTLNMNAGSIVFLTVAAADRYFKVVPVPHAVNTISTRVAAGIVCTLW 120

Qy 141 ALVILGTVYLLLENHLCVQETAVSCSFIMESANGWHDIMFQLEFFPLGLIILCSEKIV 200
Db 121 ALVILGTVYLLLENHLCVQETAVSCSFIMESANGWHDIMFQLEFFPLGLIILCSEKIV 180

Qy 201 WSLRRQLARQARMKATRIMVAIVFICVLPVPSVARSARYFLTWTPSSACDPSVHGAL 260
Db 181 WSLRRQLARQARMKATRIMVAIVFICVLPVPSVARSARYFLTWTPSSACDPSVHGAL 240

Qy 261 HTLSFTYMSMLDPLVYFSSFPKFKYKIKICSLKPKQPGHKSQTQRPPEMIS 316
Db 241 HTLSFTYMSMLDPLVYFSSFPKFKYKIKICSLKPKQPGHKSQTQRPPEMIS 296

RESULT 7
ID AAW94654 standard; Protein; 363 AA.
XX AAW94654;
XX AAW94654;
DT 29-APR-1999 (first entry)
XX G-protein coupled receptor HM74A protein.
XX HM74A receptor; G-protein coupled receptor; infection; pain; cancer;
KW diabetes; obesity; neurological disorder; heart failure; hypertension;
KW asthma; allergy.
XX Homo sapiens.
XX WO9856820-A1.
XX 17-DEC-1998.
XX 12-JUN-1998; 98WO-US12386.
XX 12-JUN-1997; 97US-0049480.
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX Bergsma DJ, Elshourbagy NA, Guerrero SF, Li X, Mooney JL;
PI WPI: 1999-095273/08.
DR N-PSDB; AAX16671.
XX New isolated G-protein coupled receptor, HM74A - used to develop
PT products for treating e.g. infections, pain, cancers, diabetes,
PT obesity, neurological disorders, heart failure, hypertension, asthma
PT or allergies
XX Claim 1; Page 31-32; 40pp; English.
XX The present sequence is a member of the G-protein coupled receptor
CC (7TM receptor) family, designated the HM74A receptor. The proteins,
CC agonists, antagonists and polynucleotides can be used for treating
CC disorders associated with increased or reduced expression or activity
CC of HM74A, e.g. bacterial, fungal, protozoan and viral infections,
CC particularly infections caused by HIV-1 or HIV-2, pain, cancers,
CC diabetes, obesity, anorexia, bulimia, asthma, Parkinson's disease,
CC acute heart failure, hypotension, hypertension, urinary retention,
CC osteoporosis, angina pectoris, myocardial infarction, stroke, ulcers,
CC asthma, allergies, benign prostatic hypertrophy, migraine, vomiting,
CC psychotic and neurological disorders, including anxiety, schizophrenia,

CC manic depression, depression, delirium, dementia, and severe mental
CC retardation, and dyskinesias such as Huntington's disease or
CC Gilles de la Tourette's syndrome. The products can also be used for
CC detection, diagnosis and drug screening.
XX
SQ Sequence 363 AA;

Query Match 47.7%; Score 883.5; DB 20; Length 363;
Best Local Similarity 52.8%; Pred. No. 1.5e-90;
Matches 180; Conservative 47; Mismatches 107; Indels 7; Gaps 4;

Qy 5 SCRIEGDITTSQVMPPLLIIVAFVGLGNGVLCGFCGFMKTPSTVYLFNLAADFLM 64
Db 17 NCCVFRDDFIVKVPVPLVGLFIFGLGNGLALMIFCFHLKSKSSRIFLNLAADFLM 76

Qy 65 MICLPFTDYLLRRHAFGDIICRVGLFTLNMNAGSIVFLTVAAADRYFKVVPVPHAV 124
Db 77 IICLPFLMDNVYRWDMKFGDIICRVGLFTLNMNAGSIVFLTVAAADRYFKVVPVPHAL 136

Qy 125 NTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCSFIMESANGWHDIMFOLE 184
Db 137 NKISNRTAAITISCLLWGITITGLTVHLLKKMPIONGGANLCSFISICHTFQWHEAMFLE 196

Qy 185 FFHPLGLIILCSEKIVWSLRRQLARQARMKATRIMVAIVFICVLPVPSVARSARYFL 244
Db 197 FFPLGLIILCSEKIVWSLRRQLARQARMKATRIMVAIVFICVLPVPSVARSARYFL 255

Qy 245 WTVPSA---CD--PSVHGALHTLSFTYMSMLDPLVYFSSFPKFKYKIKICSLK 299
Db 256 WLLHTSGTQNCVYRVDLAFITLSFTYMSMLDPLVYFSSFPKFKYKIKICSLK 315

Qy 300 KQPGHKSQTQRPPEMISNLGRRSCISVANSFQSDGQWDP 340
Db 316 KMTGPDNNRSTSVELTGDPNKT-RGAPEALMANSGEWPSP 355

RESULT 8
AAU04379
ID AAU04379 standard; Protein; 363 AA.
XX AAU04379;
XX AAU04379;
DT 23-OCT-2001 (first entry)
XX Human G-protein coupled receptor, hrUP25.
XX Human; G-protein coupled receptor; GPCR; hrUP25; agonist;
KW inverse agonist; lung cancer.
XX Homo sapiens.
XX WO200136471-A2.
XX 25-MAY-2001.
XX 16-NOV-2000; 2000WO-US31509.
XX 17-NOV-1999; 99US-0166088.
XX 17-NOV-1999; 99US-0166099.
XX 17-NOV-1999; 99US-0166369.
XX 23-DEC-1999; 99US-0171900.
XX 23-DEC-1999; 99US-0171901.
XX 23-DEC-1999; 99US-0171902.
XX 11-FEB-2000; 2000US-0181749.
XX 14-MAR-2000; 2000US-0189258.
XX 14-MAR-2000; 2000US-0189259.
XX 10-APR-2000; 2000US-0195898.
XX 10-APR-2000; 2000US-0195899.
XX 10-APR-2000; 2000US-0196078.
XX 28-APR-2000; 2000US-0200419.
XX 12-MAY-2000; 2000US-0203630.
XX 12-JUN-2000; 2000US-0210741.
XX 12-JUN-2000; 2000US-0210982.


```

Db 256 WLLHTSGTQNCVYRVDLAFFITLSTYMNMLDPVYVYFSSFPNFSTLINRCLQR 315
QY 300 KOPCHSKTORPEEMPISNLGRSCISVANSFQSDGQWDP 340
Db 316 KMTGPDNNRSTSVELTGDPNKT-RGAPEALMANSGEPWSP 355

RESULT 10
AA090637
ID AAY90637 standard; Protein; 387 AA.
AC AAY90637;
XX
XX
XX 21-AUG-2000 (first entry)
XX Human G protein-coupled receptor HM74.
XX
XX G protein-coupled receptor; GPCR; constitutively active;
KW intracellular loop 3; transmembrane domain 6; drug screening;
KW agonist; antagonist.
XX
XX Homo sapiens.
OS
XX W0200022129-AL.
XX
XX 20-APR-2000.
XX
XX 12-OCT-1999; 99WO-US23938.
XX
XX 13-OCT-1998; 98US-0170496.
XX
XX (AREN-) ARENA PHARM INC.
XX
XX Behan DP, Chalmers DT, Liaw CW;
XX
XX WPI; 2000-329165/28.
XX N-PSDB; AAA30659.
XX
XX Non-endogenous constitutively activated human G protein-coupled
PT receptors, useful for identifying agonists for use as pharmaceutical
PT agents -
XX
XX Example 1; Page 185-187; 341pp; English.
XX
XX The invention relates to constitutively active, non-endogenous versions
XX of endogenous human orphan G protein-coupled receptors (GPCRs, AAY90643-
XX AAY90677, and AAY90683-Y90687), and to DNA encoding them (AAA30709-A30743
XX and AAA30773-A30779). The mutant proteins of the invention contain a
XX mutation in a portion of the protein comprising intracellular loop 3
XX (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X,
XX is substituted for an endogenous residue in IC3 at a position 16 amino
XX acids N-terminal of an endogenous proline in TM6 to form a sequence
XX X-(AA)15-Pro. The endogenous amino acid is selected from Lys, His, Arg
XX or Ala, and is preferably Lys. When the endogenous residue at this
XX position is Lys, this residue is replaced by His, Arg or preferably Ala.
XX The 15 amino acid stretch between the substituted amino acid and the Pro
XX may be endogenous, non-endogenous, or a mixture of endogenous and
XX non-endogenous residues. The constitutively active GPCRs are useful for
XX identifying antagonists, agonists and partial agonists for use as
XX pharmaceutical agents. The mutant proteins are also useful in research
XX settings for elucidating the roles of the receptors in normal and
XX diseased conditions. Antagonists for a particular GPCR are useful for
XX treating diseases and disorders associated with that receptor. Because
XX the novel mutant GPCRs are constitutively active, they can be used
XX directly for screening of compounds without the need for endogenous
XX ligands. The present sequence represents a human wild-type GPCR referred
XX to in an exemplification of the invention.
XX
XX Sequence 387 AA;
XX
XX Query Match 47.5%; Score 880.5; DB 21; Length 387;
XX Best Local Similarity 52.2%; Pred. No. 3.6e-90;
XX
```

```

Matches 178; Conservative 49; Mismatches 107; Indels 7; Gaps 4;
QY 5 SCCRIEGDTISQVMPPLLIIVAFVGLGALNGVACGFCFHMKTWKPSVYLFNLAVADFL 64
Db 17 NCCVFRDDFAKVLPPVGLGELFGLGNGLALWIFCFHLKSMKSSRIFLNLAVADFL 76
QY 65 MICLPFRDYLLRRHWAFGDI PCRVGLFTLAMNRAGSIVFLTVVAADRYFKVYVHPH 124
Db 77 IICLPFYMDYVYRRSDWNFGDIPCRVLFLFMFAMNRQSGIIFLTVAVDYFRVYVHPH 136
QY 125 NTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCSEFIMESANGWHDIMFO 184
Db 137 NKISNMTAAIISCILWGITVGLTVHLLKLLKLLQNGPANVCISFSICHTTHERMEAME 196
QY 185 FEMPLGIILFCFSKIVMSLRROQLARQARKKATRIMVVAIVFICYLPSPVSARLYFL 244
Db 197 FLPLGLIILFCSARIWSLQR-QMDRHAKIKRAITIMVVAIVFICFLPSVVRIRIF 255
QY 245 WTVPSSA---CD--PSVHGALHITLSTYMNMLDPVYVYFSSFPFKYNNKLKICSLKP 299
Db 256 WLLHTSGTQNCVYRVDLAFFITLSTYMNMLDPVYVYFSSFPNFSTLINRCLQR 315
QY 300 KOPCHSKTORPEEMPISNLGRSCISVANSFQSDGQWDP 340
Db 316 KMTGPDNNRSTSVELTGDPNKT-RGAPEALMANSGEPWSP 355

RESULT 11
AAE02493
ID AAE02493 standard; Protein; 384 AA.
XX
XX AAE02493;
XX
XX 10-AUG-2001 (first entry)
XX
XX Human CON103 G protein-coupled receptor protein.
XX
XX Human; G protein-coupled receptor; GPCR; CON103 protein; schizophrenia;
KW neuroleptic; nootropic; neuroprotective; bipolar disease; psychotropic;
KW neurological disorder; psychiatric disease; neurosis; anxiety; neuritis;
KW attention deficit hyperactivity disorder; neuroasthenia; senile dementia;
KW affective disorder; neuropathy; Alzheimer's disease; Parkinson's disease;
KW depression; migraine; genetic screening; chromosome 2.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Domain 54..77 /label= Transmembrane_domain_(1TM)
XX Domain 78..88 /label= Intracellular_domain
XX /note= "First IC loop"
XX Domain 89..108 /label= Transmembrane_domain_(2TM)
XX /label= Transmembrane_domain_(2TM)
XX Domain 109..133 /label= Extracellular_domain
XX /label= Extracellular_domain
XX /note= "First EC loop"
XX Domain 134..149 /label= Transmembrane_domain_(3TM)
XX /label= Transmembrane_domain_(3TM)
XX Domain 150..166 /label= Intracellular_domain
XX /label= Intracellular_domain
XX /note= "Second IC loop"
XX Domain 167..188 /label= Transmembrane_domain_(4TM)
XX /label= Transmembrane_domain_(4TM)
XX Domain 189..215 /label= Extracellular_domain
XX /label= Extracellular_domain
XX /note= "Second EC loop"
XX Domain 216..240 /label= Transmembrane_domain_(5TM)
XX /label= Transmembrane_domain_(5TM)
XX Domain 241..257 /label= Intracellular_domain
XX /label= Intracellular_domain
XX /note= "Third IC loop"
XX Domain 258..283
```



```
XX SQ Sequence 423 AA;
Query Match 28.5%; Score 529; DB 20; Length 423;
Best Local Similarity 39.2%; Pred. No. 1.4e-50;
Matches 115; Conservative 48; Mismatches 108; Indels 22; Gaps 7;

Qy 4 GSCRIEDTISQVMPPLIIAVFVLGALNGVALCGFCFHKMTKPTSVYLFNLAVADFL 63
Db 81 GPCHPTSSSLVSAPLALALEFVLGLVGNLSALFIFCIHTRPMTSNTVFLVSLVAADFL 140

Qy 64 LMICLPFTDYLLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVAAADRYFKVVPHPHA 123
Db 141 LISNPLRDVYLLHETWRFGAAACKVNLMLSTNRTASVVFELTAIALNRKLVQPHHV 200

Qy 124 VNTISTRVAAGICTLWALVILGTVYLLENHLCVQE-TAVSCSFIM-----ESANGHWD 178
Db 201 LSRASVGAARVAGLWGI-----LLNGHLLSTESGSPCLSYRVGTRKPSASLRWHQ 254

Qy 179 IMFOLEFPMPGLGIILFCFSKIVWSLRRRQOLARQARKKATRFIMVVAIVTICVLPVS- 237
Db 255 ALYLLEFPLALILFAIVSGLTIRNR-GLGGQAGPORAMRVLAMVAVYTICFLPSII 313

Qy 238 --SARLYFLMTVPSSA----CDPSVHGALHITLSFTYMNMLDPLVYFSSPSF 285
Db 314 FGMASWAFWLSACRSLDLCQLPHG-----SLAFTYLSNVDPLVLYCSPSNF 362

RESULT 13
AAG78785
ID AAG78785 standard; Protein; 423 AA.
XX AC AAG78785;
XX DT 10-JAN-2002 (first entry)
XX DE Human opitoid-type receptor 1.
XX KW Human; G-protein coupled receptor; opitoid-type receptor 1; OTRL;
XX KW neurotransmission.
XX OS Homo sapiens.
XX PN DE1021475-A1.
XX PD 08-NOV-2001.
XX PF 03-MAY-2000; 2000DE-1021475.
XX PR 03-MAY-2000; 2000DE-1021475.
XX PA (BRUE/) BRUESS M.
XX PA (BOEN/) BOENISCH H.
XX PI Bruess M, Boenisch H;
XX DR WPI; 2001-657613/76.
XX DR N-PSDB; AAI71608, AAI71609.
XX PT New opitoid-type receptor-1 gene, OTRL, useful for diagnosis and
XX PT treatment of OTRL-related diseases
XX PS Disclosure; Page 5; 6pp; German.
XX CC The present invention provides the protein, cDNA and genomic DNA
XX CC sequences of human G-protein coupled receptor opitoid-type receptor 1
XX CC (OTRL). The protein is involved in neurotransmission, and the sequences
XX CC can be used in the treatment of OTRL related diseases. The present
XX CC sequence is the OTRL protein.
XX SQ Sequence 423 AA;
Query Match 28.5%; Score 529; DB 22; Length 423;
```

```
Best Local Similarity 39.2%; Pred. No. 1.4e-50;
Matches 115; Conservative 48; Mismatches 108; Indels 22; Gaps 7;

Qy 4 GSCRIEDTISQVMPPLIIAVFVLGALNGVALCGFCFHKMTKPTSVYLFNLAVADFL 63
Db 81 GPCHPTSSSLVSAPLALALEFVLGLVGNLSALFIFCIHTRPMTSNTVFLVSLVAADFL 140

Qy 64 LMICLPFTDYLLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVAAADRYFKVVPHPHA 123
Db 141 LISNPLRDVYLLHETWRFGAAACKVNLMLSTNRTASVVFELTAIALNRKLVQPHHV 200

Qy 124 VNTISTRVAAGICTLWALVILGTVYLLENHLCVQE-TAVSCSFIM-----ESANGHWD 178
Db 201 LSRASVGAARVAGLWGI-----LLNGHLLSTESGSPCLSYRVGTRKPSASLRWHQ 254

Qy 179 IMFOLEFPMPGLGIILFCFSKIVWSLRRRQOLARQARKKATRFIMVVAIVTICVLPVS- 237
Db 255 ALYLLEFPLALILFAIVSGLTIRNR-GLGGQAGPORAMRVLAMVAVYTICFLPSII 313

Qy 238 --SARLYFLMTVPSSA----CDPSVHGALHITLSFTYMNMLDPLVYFSSPSF 285
Db 314 FGMASWAFWLSACRSLDLCQLPHG-----SLAFTYLSNVDPLVLYCSPSNF 362

RESULT 14
AAU04365
ID AAU04365 standard; Protein; 423 AA.
XX AC AAU04365;
XX DT 23-OCT-2001 (first entry)
XX DE Human G-protein coupled receptor, hrup11.
XX KW Human; G-protein coupled receptor; GPCR; hrup11; agonist;
XX KW inverse agonist; lung cancer.
XX OS Homo sapiens.
XX PN WO200136471-A2.
XX PD 25-MAY-2001.
XX PF 16-NOV-2000; 2000WO-US31509.
XX PR 17-NOV-1999; 99US-0166088.
XX PR 17-NOV-1999; 99US-0166099.
XX PR 17-NOV-1999; 99US-0166369.
XX PR 23-DEC-1999; 99US-0171900.
XX PR 23-DEC-1999; 99US-0171901.
XX PR 23-DEC-1999; 99US-0171902.
XX PR 11-FEB-2000; 2000US-0181749.
XX PR 14-MAR-2000; 2000US-0189258.
XX PR 14-MAR-2000; 2000US-0189259.
XX PR 10-APR-2000; 2000US-0195898.
XX PR 10-APR-2000; 2000US-0195899.
XX PR 28-APR-2000; 2000US-0196078.
XX PR 28-APR-2000; 2000US-0200419.
XX PR 12-MAY-2000; 2000US-0203630.
XX PR 12-JUN-2000; 2000US-0210741.
XX PR 12-JUN-2000; 2000US-0210982.
XX PR 21-AUG-2000; 2000US-0226760.
XX PR 26-SEP-2000; 2000US-0235418.
XX PR 26-SEP-2000; 2000US-0235779.
XX PR 20-OCT-2000; 2000US-0242332.
XX PR 20-OCT-2000; 2000US-0242343.
XX PA (AREN-) ARENA PHARM INC.
XX PI Chen R, Dang HT, Lowitz KP;
XX WPI; 2001-355616/37.
XX DR N-PSDB; AAS07938.
XX DR
```



```
XX Endogenous and non-endogenous versions of human G-protein coupled
PT receptors for direct identification of candidate compounds as agonists,
PT inverse agonists or partial agonists for use as therapeutic agents -
XX
XX Claim 13; Page 94-96; 160pp; English.
XX
XX The sequence represents a human G-protein coupled receptor (GPCR),
XX hRUP11. The endogenous and non-endogenous, constitutively activated
CC versions of human G-protein coupled receptors (GPCR), are useful for
CC direct identification of candidate compounds as receptor agonists,
CC inverse agonists or partial agonists having applicability as therapeutic
CC agents for treating diseases related to GPCR, e.g. lung cancer.
CC Non-endogenous version of human GPCRs are also utilized in research
CC settings and in vitro and in vivo system, incorporating GPCRs can be
CC utilised to elucidate and understand the roles these receptors
CC play in the human condition, both normal and diseased.
XX
SQ Sequence 423 AA;
Query Match 28.5%; Score 529; DB 22; Length 423;
Rest Local Similarity 39.2%; Pred. No. 1.4e-50;
Matches 115; Conservative 48; Mismatches 108; Indels 22; Gaps 7;
OY 4 GSCCRIBGDTISQMPPLLIIVAFVYLGALNGVLCGFCFHMKTWKPSTVYLFNLVADFL 63
Db 81 GPCHPTSSLSVSAFLAPILALEFVLGLVGNLSALFICFIHTRPWTSTNTVFLSVLVAADFL 140
OY 64 LMTCLPRTDYLLRRHWAGDIPCRVGLFTLAMNAGSIVFLTVVAADRYFKVVPHPHA 123
Db 141 LISNLPURVDYLLHETWREGAACKVNLFMSTNRTASVVFLETAALNRYLKVYQPHV 200
OY 124 VNTISTRVAAGIVCTMALVILGVYVLLLENHLCVQE-TAVSCSEFTM---ESANGWHD 178
Db 201 LSRASVGAARVAGGLWGI-----LLNGHLLSTFGPSCLSYRVGKPKSASLRHQ 254
OY 179 IMFQLEFFPLGIILFCFKIVMSLRRLRQQLARQARKMKATRFIMVAIVFICYLPSV- 237
Db 255 ALVLEFFPLALILFAIVSIGLTRNR-GLGGQAGPQARMVLMVAVVYTCIFLPSII 313
OY 238 --SARLYFLNTVPSSA---CDPSVHGALHITLSTYNNMSMLDPLVYFFSPSPF 285
Db 314 FGMASVAFVLSACRSIDLCTQLFHG----SLAFTYLSNLDPLVLYCFSPNF 362
RESULT 15
AA94339
ID AA94339 standard; Protein; 455 AA.
XX
XX AA94339;
XX
XX 22-AUG-2000 (first entry)
XX
XX Human cell surface receptor protein #6.
XX
XX Human; HCSR; cytostatic; antiarthritic; antirheumatic; antiasthmatic;
XX immunosuppressive; antiarteriosclerotic; antibacterial; antiparasitic;
KW neuroprotective; nootropic; anticonvulsant; cancer; leukaemia;
KW melanoma; rheumatoid arthritis; asthma; atherosclerosis; akathesia;
KW Alzheimer's diseases; multiple sclerosis; epilepsy.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Domain 128..150
FT /label= Transmembrane_domain
FT Domain 291..310
FT /label= Transmembrane_domain
FT Domain 141..387
FT /label= Rhodopsin_GPCR_domain
FT Region 126..150
FT /note= "Rhodopsin signature"
FT Region 159..180
```

```
FT Region /note= "Rhodopsin signature"
FT 204..226
FT /note= "Rhodopsin signature"
FT 325..349
FT /note= "Rhodopsin signature"
FT 369..395
FT /note= "Rhodopsin signature"
FT 22
FT /note= "potential phosphorylation site"
FT 53
FT /note= "potential phosphorylation site"
FT 76
FT /note= "potential glycosylation site"
FT 189
FT /note= "potential phosphorylation site"
FT 206
FT /note= "potential phosphorylation site"
FT 207
FT /note= "potential glycosylation site"
FT 271
FT /note= "potential phosphorylation site"
FT 281
FT /note= "potential phosphorylation site"
FT 310
FT /note= "potential phosphorylation site"
FT 380
FT /note= "potential phosphorylation site"
FT 430
FT /note= "potential phosphorylation site"
XX
XX WO200028032-A2.
XX
XX 18-MAY-2000.
XX
XX 12-NOV-1999; 99NO-US26742.
XX
XX 12-NOV-1998; 98US-0191280.
XX 07-DEC-1998; 98US-0206647.
XX 08-MAR-1999; 99US-0123404.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Tang YT, Corley NC, Guegler KJ, Yue H, Baughn MR, Lal P;
XX Hillman JL, Bandman O, Azimzai Y, Au-Young J;
XX
XX WPI; 2000-376546/32.
XX N-PSDB; AAA27049.
XX
XX New human cell surface receptor protein and polynucleotide useful for
XX diagnosis, prevention and treatment of cancer, immune disorders,
XX infection and neuronal disorders -
XX
XX Claim 1; Page 79-80; 97pp; English.
XX
XX The present sequence is a novel human cell surface receptor protein
XX (HCSR) designated HCSR-6. The nucleotide sequence was identified in
XX Incyte Clone 2851578 from the cDNA library BRSTUT13, which was made
XX from RNA isolated from breast tumour tissue. A number of Incyte Clones
XX were used to assemble the consensus sequence. BLAST analysis showed that
XX the sequence is homologous to rhodopsin-like GPCR/HM74 g219867 isoform
XX g507827. HCSR and its antagonist are useful for preventing or
XX treating disorders associated with decreased or increased expression or
XX activity of HCSR. Such disorders include cancers such as leukaemia and
XX melanoma, immune disorders such as rheumatoid arthritis, asthma and
XX atherosclerosis, bacterial and parasitic infections and neuronal
XX disorders such as akathesia, Alzheimer's disease, multiple sclerosis and
XX epilepsy. Polynucleotides encoding HCSRs may be used as hybridisation
XX probes to diagnose these conditions. Anti-HCSR antibodies may be used
XX as antagonists, as a targeting or delivery mechanism for bringing
XX pharmaceutical agents into contact with cells or tissues expressing
XX HCSR and for diagnosis of HCSR-related disorders. HCSR and its
XX catalytic or immunogenic fragments are useful for drug screening using
XX libraries of compounds.
```

```

XX  SQ  Sequence  455 AA;
Query Match      28.5%; Score 529; DB 21; Length 455;
Best Local Similarity 39.2%; Pred. No. 1.5e-50;
Matches 115; Conservative 48; Mismatches 108; Indels 22; Gaps 7;

Qy  4  GSCCRIEGDTISQVMPPLLIYAFVLGALNGCVALCGCFHMKTKWKPSTVYLFNLAVADFL 63
Db  113  GPCHPTSSSLVSAPLALALEFVLGLVGNLSLAFICIHTRPWTSTNTVFLVSLVAADFL 172

Qy  64  LMICLPFTDYLLRRRHWFAGDIPCRVGLFTLANNRAGSIVFLTVVAADRYKVVHPHHA 123
Db  173  LISNPLRVVDYLLHETWRFGAAACKVNLFMLSTNRTASVVFLTAIALNRYLKVVQPHHV 232

Qy  124  VNTISTRVAAGIVCTLWALVILGVYLLLENHLCVQE-TAVSCBSFTM-ESANGWHD 178
Db  233  LSRASVGAAARVAGGLWGI-----LLNGHLLSTFGSPCLSYRVGTRKPSASLRWHQ 286

Qy  179  IMFQLEFFMPLGIILFCFSKIVWSLRRRQQLARQARKKATRFIMVVAIVITCYLPSV- 237
Db  287  ALYLLEFPLALILFAIVSIGLIRNR-GLGGQAGPQARNRVLAMVVAVYTICFLPSII 345

Qy  238  --SARLYFLWTVPSSA---CDPSVHGALHITLSETYMNMSMLDPLVYFSPSPF 285
Db  346  FGMASWVAFWLSACRSLDLCTQLFHG-----SLAFTYLSNVLDPLVLYCFSPNF 394
```

Search completed: October 29, 2002, 04:52:11
Job time : 64 secs

PR 16-NOV-1999; 99US-0165838.
 PR 17-NOV-1999; 99US-0166071.
 PR 19-NOV-1999; 99US-0166678.
 PR 28-DEC-1999; 99US-0173396.
 PR 22-FEB-2000; 2000US-0184129.
 PR 28-FEB-2000; 2000US-0185421.
 PR 28-FEB-2000; 2000US-0185554.
 PR 02-MAR-2000; 2000US-0186530.
 PR 03-MAR-2000; 2000US-0186811.
 PR 09-MAR-2000; 2000US-0188114.
 PR 17-MAR-2000; 2000US-0190310.
 PR 21-MAR-2000; 2000US-0190800.
 PR 20-APR-2000; 2000US-0198568.
 PR 02-MAY-2000; 2000US-0201190.
 PR 08-MAY-2000; 2000US-0203111.
 PR 25-MAY-2000; 2000US-0207094.
 XX
 PA (PHAA) PHARMACIA & UPJOHN CO.
 XX
 XX Vogeli G, Wood LS, Parodi LA, Hiebsch RB, Lind P, Slightom J;
 PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;
 XX
 DR WPI; 2001-389826/41.
 DR P-PSDB; AAG80968.
 XX
 PT New G protein-coupled receptor (ngPCR-x) and its encoding
 PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -
 XX
 PS Claim 4; Page 89; 261pp; English.
 XX
 CC The present invention relates to novel G protein-coupled receptors
 CC (ngPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
 CC 28, 31-38, 40, 41, 53-60) and their coding sequences. The present
 CC sequence is the coding sequence for one such G protein-coupled receptor.
 CC GPCR are also known as seven transmembrane receptors and function in
 CC signal transduction. The ngPCRx coding sequences are useful for
 CC screening a human to diagnose a disorder affecting the brain or a genetic
 CC predisposition, specifically schizophrenia. ngPCRx are useful for
 CC identifying compounds useful for treating schizophrenia. Detection of
 CC ngPCRx in a sample is useful as a diagnostic tool for diseases or
 CC disorders e.g. thyroid disorders, renal failure, rheumatoid arthritis,
 CC CNS disorders, infections such as HIV-1, metabolic and cardiovascular
 CC diseases, proliferative disorders and hormonal disorders. Modulators of
 CC ngPCRx activity have the utility for treating neurological disorders,
 CC including schizophrenia, ADHD/ADD (attention deficit-hyperactivity
 CC disorder/attention deficit disorder), and neuronal disorders such as
 CC Alzheimer's disease, Parkinson's disease, migraine and senile dementia.
 CC Additional disorders include inflammatory conditions (e.g. Crohn's
 CC disease), rheumatoid arthritis, autoimmune disorders, cancers,
 CC respiratory ailments such as asthma, and inflammatory diseases e.g.
 CC inflammatory bowel disease.
 XX
 SQ Sequence 1041 BP; 208 A; 294 C; 278 G; 261 T; 0 other;
 Query Match 100.0%; Score 1041; DB 22; Length 1041;
 Best Local Similarity 100.0%; Pred. No. 4e-288;
 Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Yy 1 ATGTACACGGGTGCTGTCGGCGCATCGAGGGGACACCATCTCCAGGTGATGCCGCG 60
 Db 1 ATGTACACGGGTGCTGTCGGCGCATCGAGGGGACACCATCTCCAGGTGATGCCGCG 60
 Yy 61 CTGCTCATTTGGCCCTTTGCTGCGGCGCACTAGGCAATGGGTGCGCCCTGTGGTTTC 120
 Db 61 CTGCTCATTTGGCCCTTTGCTGCGGCGCACTAGGCAATGGGTGCGCCCTGTGGTTTC 120
 Yy 121 TGCTTCCACATGAAGACCTTGAAGCCCGACGACTGTTTACCTTTCAATTGGCCGTGCT 180
 Db 121 TGCTTCCACATGAAGACCTTGAAGCCCGACGACTGTTTACCTTTCAATTGGCCGTGCT 180
 Yy 181 GATTTCCTCTTATGATCTGCTGCTTTTGGGACAGACTATTACCTCAGACGATAGACAC 240
 Db 181 GATTTCCTCTTATGATCTGCTGCTTTTGGGACAGACTATTACCTCAGACGATAGACAC 240

Qy 241 TGGGCTTTTGGGACATTCCTCGCGAGTGGGGCTCTTACAGTTGGCCATGAACAGGCC 300
 Db 241 TGGGCTTTTGGGACATTCCTCGCGAGTGGGGCTCTTACAGTTGGCCATGAACAGGCC 300
 Qy 301 GGGAGCATCGTGTTCCTTACGCTGGTGGGACAGGTATTTCAAAGTGGTCCACCCC 360
 Db 301 GGGAGCATCGTGTTCCTTACGCTGGTGGGACAGGTATTTCAAAGTGGTCCACCCC 360
 Qy 361 CACCACGGGTGAACACTATCTCCACCCGGGTGGCGGTGGCATCTGTGCACCCCTGTGG 420
 Db 361 CACCACGGGTGAACACTATCTCCACCCGGGTGGCGGTGGCATCTGTGCACCCCTGTGG 420
 Qy 421 GCCCTGGTCACTCCCTGGGAAACAGTGTATCTTTTGTGGAGAACCATCTCTGGCTGCAAG 480
 Db 421 GCCCTGGTCACTCCCTGGGAAACAGTGTATCTTTTGTGGAGAACCATCTCTGGCTGCAAG 480
 Qy 481 ACGGCGCTCTCTGTGAGAGCTTCATCATGAGTCGGCAATGGCTGGCATGACATCATG 540
 Db 481 ACGGCGCTCTCTGTGAGAGCTTCATCATGAGTCGGCAATGGCTGGCATGACATCATG 540
 Qy 541 TTCCAGCTGGAGTTCCTTATGCCCTCGGCATCATCTTATTTTGTCTCTCAAGATTGT 600
 Db 541 TTCCAGCTGGAGTTCCTTATGCCCTCGGCATCATCTTATTTTGTCTCTCAAGATTGT 600
 Qy 601 TGGAGCTTGAGCGGAGGACAGCTGCCAGACAGGCTCGGATGAAGAGCGACCCGG 660
 Db 601 TGGAGCTTGAGCGGAGGACAGCTGCCAGACAGGCTCGGATGAAGAGCGACCCGG 660
 Qy 661 TTCATCATGGTGGTGGCAATTTGTTTCACTACATGCTACCTGCCAGCGTGTCTCTAGA 720
 Db 661 TTCATCATGGTGGTGGCAATTTGTTTCACTACATGCTACCTGCCAGCGTGTCTCTAGA 720
 Qy 721 CTCATTTTCTCTGAGCGGTGCCCTCGAGTGCCTCGCATCCCTCTGTCCATGGGCGCTG 780
 Db 721 CTCATTTTCTCTGAGCGGTGCCCTCGAGTGCCTCGCATCCCTCTGTCCATGGGCGCTG 780
 Qy 781 CACATAACCTTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTCGGTGTTATTTT 840
 Db 781 CACATAACCTTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTCGGTGTTATTTT 840
 Qy 841 TCAAGCCCTCTCTTCCCAATTTCTACAAAGCTCAAAATCTGCAGTCTGAACCCCAAG 900
 Db 841 TCAAGCCCTCTCTTCCCAATTTCTACAAAGCTCAAAATCTGCAGTCTGAACCCCAAG 900
 Qy 901 CAGCAGGACACTCAAAACAAAGGCGGGAAGAGATGCCAATTTTGAACCTCGGTGCG 960
 Db 901 CAGCAGGACACTCAAAACAAAGGCGGGAAGAGATGCCAATTTTGAACCTCGGTGCG 960
 Qy 961 AGGAGTTCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGTATGGCAATGGGATCCC 1020
 Db 961 AGGAGTTCATCAGTGTGGCAATAGTTTCCAAAGCCAGTCTGTATGGCAATGGGATCCC 1020
 Qy 1021 CACATTTGTGAGTGGCACTGA 1041
 Db 1021 CACATTTGTGAGTGGCACTGA 1041

RESULT 4
 ABA81529
 ID ABA81529 standard; DNA; 1050 BP.
 XX ABA81529;
 XX
 XX
 DT 28-JAN-2002 (first entry).
 XX
 DE Human GPCR1a polynucleotide SEQ ID NO 1.
 XX
 KW Human; GPCR; G-coupled protein-receptor; cardiant; antiarteriosclerotic;
 KW anabolic; cytosstatic; antiviral; gene therapy; cardiomyopathy; obesity;
 KW anorexia; diabetes; osteoporosis; Crohn's disease; multiple sclerosis;
 KW asthma; Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KW infection; human immunodeficiency virus; HIV; ds.

THIS PAGE BLANK (U.S. . . .)

Best Local Similarity 100.0%; Pred. No. 6.8e-168; Matches 296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LLIVAFVGLGNGVAGLGCFCFHMKTWKPSTVYLFNLAVADFLLMICLPFRDYYLRRRH 80
DB 1 LLIVAFVGLGNGVAGLGCFCFHMKTWKPSTVYLFNLAVADFLLMICLPFRDYYLRRRH 60

QY 81 WAFGDIPOCVGLFTLNMNRAGSIVFLTVVADRYKVVPHHVNITSTRVAAGIVCTLM 140
DB 61 WAFGDIPOCVGLFTLNMNRAGSIVFLTVVADRYKVVPHHVNITSTRVAAGIVCTLM 120

QY 141 ALVILGTVYLLLENHLCVQETAVSCETIMESANGWHDIMFOLEFPMPLGILFCFKIV 200
DB 121 ALVILGTVYLLLENHLCVQETAVSCETIMESANGWHDIMFOLEFPMPLGILFCFKIV 180

QY 201 WSLRRQOLARQARKKATRTIMVVAIVFICYLPSVSARLYFLWTVPSACDPSVHGAL 260
DB 181 WSLRRQOLARQARKKATRTIMVVAIVFICYLPSVSARLYFLWTVPSACDPSVHGAL 240

QY 261 HITLSFTYMNMLDPLVYFSSPPKPKYKIKICSLKPKQPGHSGKTQRPPEMPTIS 316
DB 241 HITLSFTYMNMLDPLVYFSSPPKPKYKIKICSLKPKQPGHSGKTQRPPEMPTIS 296

RESULT 7
AAW94654
ID AAW94654 standard; Protein; 363 AA.
AC AAW94654;
XX
DT 29-APR-1999 (first entry)
XX
DE G-protein coupled receptor HM74A protein.
XX
KW HM74A receptor; G-protein coupled receptor; infection; pain; cancer;
KW diabetes; obesity; neurological disorder; heart failure; hypertension;
KW asthma; allergy.
XX
OS Homo sapiens.
XX
PN WO9856820-A1.
XX
PD 17-DEC-1998.
XX
PF 12-JUN-1998; 98WO-US12386.
XX
PR 12-JUN-1997; 97US-0049480.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Bergsma DJ, Elshourbagy NA, Guerrero SF, Li X, Mooney JL;
XX
DR WPI: 1999-095273/08.
XX
N-PSDB: AAX16671.
XX
PT New isolated G-protein coupled receptor, HM74A - used to develop
PT products for treating e.g. infections, pain, cancers, diabetes,
PT obesity, neurological disorders, heart failure, hypertension, asthma
PT or allergies
XX
PS Claim 1; Page 31-32; 40pp; English.
XX
CC The present sequence is a member of the G-protein coupled receptor
CC (7TM receptor) family, designated the HM74A receptor. The proteins,
CC agonists, antagonists and polynucleotides can be used for treating
CC disorders associated with increased or reduced expression or activity
CC of HM74A, e.g. bacterial, fungal, protozoan and viral infections,
CC particularly infections caused by HIV-1 or HIV-2, pain, cancers,
CC diabetes, obesity, anorexia, bulimia, asthma, Parkinson's disease,
CC acute heart failure, hypertension, hypertension, urinary retention,
CC osteoporosis, angina pectoris, myocardial infarction, stroke, ulcers,
CC asthma, allergies, benign prostatic hypertrophy, migraine, vomiting,
CC psychotic and neurological disorders, including anxiety, schizophrenia,

CC manic depression, depression, delirium, dementia, and severe mental
CC retardation, and dyskinesias such as Huntington's disease or
CC Gilles de la Tourette's syndrome. The products can also be used for
CC detection, diagnosis and drug screening.
XX
SQ Sequence 363 AA;

Query Match 47.7%; Score 883.5; DB 20; Length 363;
Best Local Similarity 52.8%; Pred. No. 1.5e-90;
Matches 180; Conservative 47; Mismatches 107; Indels 7; Gaps 4;

QY 5 SCCRIEGDTISQVMPPLIIVAFVGLGNGVAGLGCFCFHMKTWKPSTVYLFNLAVADFL 64
DB 17 NCCVFRDDFIYKVLPPVLGLEIFGLGNGLALWIFCFLKSKWSRIFLFLNLAVADFL 76

QY 65 MICLPFTDYLLRRHNAFGDIPCRVGLFTLNMNRAGSIVFLTVVADRYKVVPHHVNIT 124
DB 77 IICLPFLMDNVRWDKFGDIPCRVGLFTLNMNRAGSIVFLTVVADRYKVVPHHVNIT 136

QY 125 NTISTRVAAGIVCTLMALVILGTVYLLLENHLCVQETAVSCETIMESANGWHDIMFOLE 184
DB 137 NKISNRATAAIIISCLLWITIGLTVHLLKKMPIONGGANLCSFSICHTFQWHEAMFLE 196

QY 185 FPMPLGILFCFKIVLSRRQOLARQARKKATRTIMVVAIVFICYLPSVSARLYFL 244
DB 197 FFLPLGILFCSARIISLRQ-QMDRHAIKRAITFMVVAIVFICFLPSVYVRIRIF 255

QY 245 WTVPSA--CD--PSVHGALHITLSFTYMNMLDPLVYFSSPPKPKYKIKICSLKPK 299
DB 256 WLLHTSGTQNCVYRSVDLAFFITLSFTYMNMLDPLVYFSSPPKPKYKIKICSLKPK 315

QY 300 KQGHSGKTQRPPEMPTISNLGRRSISVANSFQSDGOWDP 340
DB 316 KWTGPDNNRSTVELATGDPNKT-RGAPEALMANSGEWSP 355

RESULT 8
AAU04379
ID AAU04379 standard; Protein; 363 AA.
XX
AC AAU04379;
XX
DT 23-OCT-2001 (first entry)
XX
DE Human G-protein coupled receptor, hRUP25.
XX
KW Human; G-protein coupled receptor; GPCR; hRUP25; agonist;
KW inverse agonist; lung cancer.
XX
OS Homo sapiens.
XX
PN WO200136471-A2.
XX
PD 25-MAY-2001.
XX
PF 16-NOV-2000; 2000WO-US31509.
XX
PR 17-NOV-1999; 99US-0166088.
XX
PR 17-NOV-1999; 99US-0166099.
XX
PR 23-DEC-1999; 99US-0171900.
XX
PR 23-DEC-1999; 99US-0171901.
XX
PR 23-DEC-1999; 99US-0171902.
XX
PR 11-FEB-2000; 2000US-0181749.
XX
PR 14-MAR-2000; 2000US-0189258.
XX
PR 10-APR-2000; 2000US-0195259.
XX
PR 10-APR-2000; 2000US-0195898.
XX
PR 10-APR-2000; 2000US-0195899.
XX
PR 28-APR-2000; 2000US-0200419.
XX
PR 12-MAY-2000; 2000US-0203630.
XX
PR 12-JUN-2000; 2000US-0210741.
XX
PR 12-JUN-2000; 2000US-0210982.

THIS PAGE BLANK (USPTO,

Oy 790 CTCAGCTTCACCTACATGAACAGCATGCTGGA 821
 Db 838 CTCAGCTTCACCTACATGAACAGCATGCTGGA 869

RESULT 14

AA16671
 ID AAX16671 standard; cDNA; 1361 BP.
 XX
 AC AAX16671;

XX 29-APR-1999 (first entry)

XX G-protein coupled receptor HM74A encoding cDNA.

XX HM74A receptor; G-protein coupled receptor; infection; pain; cancer;
 KW diabetes; obesity; neurological disorder; heart failure; hypertension;
 KW asthma; allergy; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
 PH 61.1152
 FT CDS
 FT /*tag= a
 FT /product= "HM74A receptor"

XX WO9856820-A1.

XX 17-DEC-1998.

XX 12-JUN-1998; 98WO-US12386.

XX 12-JUN-1997; 97US-0049480.

XX (SMK) SMITHKLINE BEECHAM CORP.

XX Bergsma DJ, Elshourbagy NA, Guerrero SF, Li X, Mooney JL;

XX P-PSDB; AAW94654.

XX New isolated G-protein coupled receptor, HM74A - used to develop
 PT products for treating e.g. infections, pain, cancers, diabetes,
 PT obesity, neurological disorders, heart failure, hypertension, asthma
 PT or allergies

XX Claim 2; Page 30-31; 40pp; English.

XX The present sequence encodes a member of the G-protein coupled receptor
 CC (7TM receptor) family, designated the HM74A receptor. The proteins,
 CC agonists, antagonists and polynucleotides can be used for treating
 CC disorders associated with increased or reduced expression or activity
 CC of HM74A, e.g. bacterial, fungal, protozoan and viral infections,
 CC particularly infections caused by HIV-1 or HIV-2, pain, cancers,
 CC diabetes, obesity, anorexia, bulimia, asthma, Parkinson's disease,
 CC acute heart failure, hypotension, hypertension, urinary retention,
 CC osteoporosis, angina pectoris, myocardial infarction, stroke, ulcers,
 CC asthma, allergies, benign prostatic hypertrophy, migraine, vomiting,
 CC psychotic and neurological disorders, including anxiety, schizophrenia,
 CC manic depression, depression, delirium, dementia, and severe mental
 CC retardation, and dyskinesias such as Huntington's disease or
 CC Gilles de la Tourette's syndrome. The products can also be used for
 CC detection, diagnosis and drug screening.

XX Sequence 1361 BP; 291 A; 390 C; 342 G; 338 T; 0 other;

Query Match 3.18; Score 32; DB 20; Length 1361;

Best Local Similarity 100.0%; Pred. No. 8.9e-06;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 790 CTCAGCTTCACCTACATGAACAGCATGCTGGA 821

Db 838 CTCAGCTTCACCTACATGAACAGCATGCTGGA 929

RESULT 15

AAH51053

ID AAH51053 standard; DNA; 24 BP.

XX AAH51053;

XX 28-AUG-2001 (first entry)

XX Oligonucleotide SEQ ID 133.

XX G protein-coupled receptor; nPCR; seven transmembrane receptor;
 KW signal transduction; schizophrenia; thyroid disorder; renal failure;
 KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
 KW cardiovascular disease; proliferative disorder; hormonal disorder;
 KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
 KW attention deficit-hyperactivity disorder; attention deficit disorder;
 KW Parkinson's disease; migraine; senile dementia; inflammatory disease;
 KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;
 KW neuroprotective; ss.

XX Synthetic.

XX WO200136473-A2.

XX 25-MAY-2001.

XX 16-NOV-2000; 2000WO-US31581.

XX 16-NOV-1999; 99US-0165838.

XX 17-NOV-1999; 99US-0166071.

XX 19-NOV-1999; 99US-0166678.

XX 28-DEC-1999; 99US-0173396.

XX 22-FEB-2000; 2000US-0184129.

XX 28-FEB-2000; 2000US-0185421.

XX 28-FEB-2000; 2000US-0185554.

XX 02-MAR-2000; 2000US-0186530.

XX 03-MAR-2000; 2000US-0186811.

XX 09-MAR-2000; 2000US-0188114.

XX 17-MAR-2000; 2000US-0190310.

XX 21-MAR-2000; 2000US-0190800.

XX 20-APR-2000; 2000US-0198568.

XX 02-MAY-2000; 2000US-0201190.

XX 08-MAY-2000; 2000US-0203111.

XX 25-MAY-2000; 2000US-0207094.

XX (PHAA) PHARMACIA & UPJOHN CO.

XX Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;

XX Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;

XX WPI; 2001-389826/41.

XX New G protein-coupled receptor (ngPCR-x) and its encoding
 PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -
 PS Disclosure; Page 246; 261pp; English.

XX The present invention relates to novel G protein-coupled receptors
 CC (ngPCR; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
 CC 28, 31-38, 40, 41, 53-60) and their coding sequences (see
 CC AAH50969-AAH51015 and AAH51105 and AAG80929-AAG80975 and AAG80977). The
 CC present sequence is an oligonucleotide, which was used in the present
 CC invention. GPCRs are also known as seven transmembrane receptors and
 CC function in signal transduction. The ngPCR coding sequences are useful
 CC for screening a human to diagnose a disorder affecting the brain or a
 CC genetic predisposition, specifically schizophrenia. ngPCR are useful
 CC for identifying compounds useful for treating schizophrenia. Detection of
 CC ngPCR in a sample is useful as a diagnostic tool for diseases or
 CC disorders e.g. thyroid disorders, renal failure, rheumatoid arthritis,
 CC CNS disorders, infections such as HIV-1, metabolic and cardiovascular
 CC diseases, proliferative disorders and hormonal disorders. Modulators of

THIS PAGE BLANK (USPTO)
